Lawrence H. Officer

Essays in Economic History

Purchasing Power Parity, Standard of Living, and Monetary Standards

palgrave
macmillan

Essays in Economic History

Lawrence H. Officer

Essays in Economic History

Purchasing Power Parity, Standard of Living, and Monetary Standards

palgrave
macmillan

Lawrence H. Officer
University of Illinois at Chicago
Glencoe, IL, USA

ISBN 978-3-030-95924-1 ISBN 978-3-030-95925-8 (eBook)
https://doi.org/10.1007/978-3-030-95925-8

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

In Memoriam—My Mentors
Richard E. Caves
Gottfried Haberler
Hendrik S. Houthakker

"Ignorance is the curse of God, Knowledge the wing wherewith we fly to heaven."—Shakespeare

Preface

This collection of essays is the second assemblage of my writings, but should have been the first. In the first collection (Officer, 2007), I deliberately selected a variety of subjects secondary to my main work. To right the matter, this book's collection consists of three topics on which I have spent the major part of my research: purchasing power parity, standard of living, and monetary standards. In addition, to end the collection, there is an original work "Economics and Economic History in Science Fiction." I thank my younger son, Ari J. Officer, for taking the lead in this essay and graciously accepting me as co-author. While science fiction is an avocation for both Ari and myself, it is a fair statement that contributing to that literature is closer to a true vocation for Ari.

In reviewing the first collection, Richard Sylla (2007) writes:

> "As they contemplate mortality and immortality," the late Charles Kindleberger (1985, 1) once wrote, "many economists … think it useful to gather their scattered academic detritus into packages, organized either chronologically or by subject." Kindleberger was a master of the genre, producing several such packages, which he described as exercises in tidying up things for one's literary executor. In case you hadn't guessed from the title of Lawrence Officer's new book, it is a recent addition to the genre.

That description also applies to the present volume, concerning which I thank three anonymous reviewers for their comments. Nothing is more important to an author than the respect of one's peers.

As I scan my research life, the indebtedness to many people looms large. While the current book (with exception of the science-fiction essay) consists of sole-authored items, these writings benefitted immensely from the rich experience of joint work with other scholars. Wherefore I acknowledge that I learned many things about scholarly research from co-authors, who mainly were at institutions with which I have been associated. At Harvard University: Lawrence B. Smith, Thomas D. Willett, Thomas A. Wilson; at Bank of Canada: John F. Helliwell, Harold T. Shapiro, Ian A. Stewart; at International Monetary Fund: Morris Goldstein, Mohsin S. Khan; at Michigan State University: Mordechai E. Kreinin, Daniel H. Saks, Judith A. Saks, Leanna Stiefel; at MeasuringWorth.com: Samuel H. Williamson; without joint affiliation: Marina Cristina Marcuzzo, Ari J. Officer, Annalisa Rosselli.

This book is in memoriam of my mentors at Harvard University. When I arrived there as a graduate student in 1960, I was a 20-year-old with little life experience and even less self-confidence. Professors Richard E. Caves. Gottfried Haberler, and Hendrik S. Houthakker saw potential in me that I did not see in myself and treated me as a scholar when I was far from deserving that title. And on a personal level: Houthakker persuaded me to continue in the program when I was discouraged and about to leave in the first semester, Caves guided me through the workings of the Economics Department, and Haberler invited me to many one-on-one Sunday breakfasts to discuss economics (much later, I had the opportunity to salute Gottfried Haberler's contributions to macroeconomics–Officer, 1982). Without the kindness of these three exemplary human beings, I should never have been in a position to have writings to collect for either the present or past volume!

Glencoe, IL, USA — Lawrence H. Officer
August 2022

References

Kindleberger, Charles P. (1985). *Keynesianism vs. Monetarism, and Other Essays in Financial History*. London: George Allen & Unwin.

Officer, Lawrence H. (1982). "Prosperity and Depression—And Beyond." *Quarterly Journal of Economics* 97 (February), pp. 149–159.

______. (2007). *Pricing Theory, Financing of International Organizations and Monetary History*. London and New York: Routledge.

Sylla, Richard. (2007). Review of Lawrence H. Officer, *Pricing Theory, Financing of International Organizations and Monetary History*, EH.Net.

Previous Books of Lawrence H. Officer

An Econometric Model of Canada Under the Fluctuating Exchange Rate, Harvard University Press, 1968.

The International Monetary System: Problems and Proposals (co-editor), Prentice-Hall, 1969.

Canadian Economic Problems and Policies (co-editor), McGraw-Hill, 1970.

Supply Relationships in the Canadian Economy: An Industry Comparison, International Business and Economic Studies, Michigan State University, 1972.

Issues in Canadian Economics (co-editor), McGraw-Hill Ryerson, 1974.

The Monetary Approach to the Balance of Payments: A Survey (co-author), International Finance Section, Princeton University, 1978.

So You Have to Write an Economics Term Paper… (co-author), Michigan State University Press, 1981.

Purchasing Power Parity: Theory, Evidence and Relevance, JAI, 1982.

International Economics (editor), Kluwer-Nijhoff, 1987.

Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior, 1791–1931, Cambridge University Press, 1996.

Monetary Standards and Exchange Rates (co-editor), Routledge, 1997.

Pricing Theory, Financing of International Organizations and Monetary History, Routledge, 2007.

Two Centuries of Compensation for U.S. Production Workers in Manufacturing, Palgrave Macmillan, 2009.

Everyday Economics: Honest Answers to Tough Questions, Palgrave Macmillan, 2009.

Translated into Chinese (traditional characters), Wealth, 2009.

Translated into Chinese (simplified characters), CITIC, 2010.

Translated into German, Börsenbuchverlag, 2013.

A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services, Palgrave Macmillan, 2021.

Contents

List of Figures

List of Tables

PART I

Purchasing Power Parity: Origin and Use

CHAPTER 1

Salamancans and Gerard Malynes

1.1 Salamanca School, Tudor Period

Originally published in *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance*, JAI, 1982, pp. 30–36.

1.1.1 Salamanca School

The originators of the purchasing power parity (PPP) theory were Spanish scholars of the sixteenth-century, the Salamanca School. As will be shown below, there can be no doubt about this assertion. Yet the Salamancan accomplishment went unnoticed in the English literature until Margorie Grice-Hutchinson (1952) authored a description of Spanish monetary theory in the 1544–1605 period, while providing translated excerpts from the writings of the scholars involved. Later, basing his comments on Grice-Hutchinson's work, Einzig (1970) also attributed the origins of the PPP theory to these sixteenth and seventeenth century Spanish writers.[1]

It was a confluence of diverse circumstances that led these scholars to develop the PPP theory. First of all, by the middle of the sixteenth century the University of Salamanca, in western Spain, was a great center of learning the seventy chairs of which, according to Grice-Hutchinson (1952, p. xi), were "filled by the best scholars of the age." Second, these scholars, as theologians and jurists, were well acquainted with the earlier,

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_1

scholastic work on foreign exchange. Indeed, Grice-Hutchinson views the Salamancan analysis of foreign exchange as a development of the theories of the Florentine theologians Laurentius and St. Antonio.

Yet, and third, the people of the Salamanca School could not help but be interested in secular issues, among which was international commercial activity, for which Spain had become a leading center. This role of Spain was closely related to its conquests in America and the resulting flow of gold and silver to the home country. Fourth, Medieval analysis of foreign exchange had included the idea that ease (scarcity) of a money gave it a low (high) value against foreign exchange.[2] The missing link to reach the PPP theory was the quantity theory of money.[3] The empirical impetus for the quantity theory was provided in 16th-century Spain, the first country in Europe to receive large inflows of precious metals from the New World, with resultant conspicuous increases in the money supply and in prices.

It should be noted that the true contribution of New World treasure to the sixteenth-century "price revolution" is beside the point for our purposes here. No doubt, other factors were involved, including those on the real side. The *perception* of substantial increases in the coined money supply and in prices led to the formulation of the quantity-theory relationship between the two, and earliest in Spain; that is the relevant point.

Fifth, it was also clear empirically that exchange rates had become unfavorable to Spain. If exchange rates themselves were not recorded, nevertheless, according to Grice-Hutchinson, the Salamancan economists observed that the ratio of the amount of money repaid to the amount delivered was much higher for initial delivery of money to Spain from abroad than this two-way transaction beginning in the opposite direction. This relationship required an explanation, and relative supplies of money or relative price levels in Spain and foreign countries were obvious candidates.

Sixth, premiums on exchange transactions incorporating a time element (that is, on bills of exchange) had long been used as a way of escaping the Catholic Church's prohibition of usury. The Salamancans had a theological benefit in developing a theory such as PPP; variations in exchange rates could then be interpreted as non-usurious in nature and so quite consistent with Church doctrine. Grice-Hutchinson writes: "This early version of the purchasing-power parity theory…removed the taint of usury that had formerly accompanied even the most genuine exchange transaction" (1952, p. 58). Indeed, she explains the demise of the PPP theory in the late seventeenth century as reflecting a final full toleration of exchange transactions, irrespective of their nature:

> The last traces of the medieval objection to exchange transactions (though not, of course, the dislike of usury itself) seems to have died away towards the end of the seventeenth century. . .the old purchasing-power parity theory, which had been framed to show that the premium on a bill of exchange was not necessarily a disguised form of interest on a loan, lost its *raison d'etre* and presumably died a natural death after performing a useful function for close on 150 years. (1952, p. 77)

In spite of these common circumstances, not all the Salamancan writers on exchange-rate determination put forward the PPP theory. To some extent, this may have been due to the natural development of the PPP approach from antecedent theories in an atmosphere in which the scholars had access to, and commented on, each-other's work. Another reason, no doubt, is that some Salamancans preferred to profess alternative theories of foreign exchange even while aware of the PPP approach. These other theories amounted to sophisticated treatments of the demand-and-supply and money-supply theories that developed in the Middle Ages.

Our concern here is with those Spanish writers that proposed the PPP theory itself. The earliest of these, and certainly a forerunner if not the actual founder of the PPP approach, is Azpilcueta de Navarro, writing in 1556. In any event, he is without doubt the founder of the quantity theory of money; for he writes[4]:

> other things being equal, in countries where there is a great scarcity of money, all other saleable goods, and even the hands and labour of men, are given for less money than where it is abundant... And even in Spain, in times when money was scarcer, saleable goods and labour were given for very much less than after the discovery of the Indies, which flooded the country with gold and silver. (Quoted by Grice-Hutchinson, 1978, p. 104)

The PPP theory is presented in a less direct fashion. Navarro states: "We cannot know whether an exchange transaction be just unless we know the value of both monies; since… the money must be changed at its proper value if the transaction is to be a just one." He then presents various reasons why "the value of the two moneys may diverge," among which "because of scarcity and need." Concentrating on this reason, he declares that "money, in so far as it may be sold, bartered, or exchanged by some other form of contract, is merchandise and therefore also becomes dearer when it is in great demand and short supply."[5] He then proceeds to make the connection between the scarcity or abundance of money and the high or low level of prices, via the quantity theory of money in the passage quoted above. The result is the relative-PPP theory.

The Salamancan writers are considering coined, not paper, money. When Navarro states that "the value of the two moneys may diverge," his standard of reference must be the mint parity between the monies. The "proper value" of the exchange rate is not the mint parity, but the PPP. It is PPP that explains deviations of exchange rates from mint parities.

Though Navarro thus formulates the PPP theory in an indirect fashion, it is a complete statement of the theory in that the discussions of monetary ease and scarcity and of the quantity theory are general in nature, therefore applicable to both the domestic and foreign country.

In 1594, Domingo de Bañez stated the PPP theory quite directly:

> In places where money is scarce, goods will be cheaper than in those where the whole mass of money is bigger, and therefore it is lawful to exchange a smaller sum in one country for a larger sum in another.... one party may lawfully agree to repay a larger sum to another, corresponding to the amount required to buy the same parcel of goods that the latter might have bought if he had not delivered his money in exchange." (Cited in Grice-Hutchinson, 1952, pp. 57–58)

Again, sums of money in different currencies can be compared only via some standard, implicitly the mint parity. The exchange value of a country's money can legitimately exceed its mint parity when the money's purchasing power over commodities exceeds that of money abroad. This is a theory of absolute PPP in which currencies exchange with each other in their respective amounts that are required to purchase the same basket ("parcel") of goods. A similar presentation of PPP theory was made by Juan de Lugo in 1642:

> the excess of this unequal value which money has in different places... may also be caused by diversity in its extrinsic value. Thus, in the place to which the money is sent there may be a general scarcity of money, or more people may require it, or there may be better opportunities for doing business with it and making a profit. And, since money will there be more useful for satisfying human needs, more goods will be bought than elsewhere with the same sum of money, and therefore money will rightly

> be regarded as more valuable in that place. (Quoted by Grice- Hutchinson 1978, p. 106)

Once more, a Salamancan scholar is presenting the absolute-PPP theory. The exchange rate between two currencies, expressed as a deviation from their metallic parity ("the excess of this unequal value which money has in different places"), is determined by the relative purchasing power ("extrinsic value") of the monies. (In the unquoted part of the passage, de Lugo points out that another determinant of the exchange rate is differing "intrinsic value," metallic content, of monies.)

At first consideration, it seems surprising that the PPP theory was developed not under a freely floating exchange rate, with unconstrained exchange-rate movements, but rather under a metallic standard, with the exchange rate confined within specie points. Yet, to repeat, the latter situation applied. Gold and silver coins (or bills of exchange payable in coin) were the usual medium of foreign-exchange transactions. An unconstrained floating rate for Spain would have involved paper money irredeemable in gold or silver.

However, upper and lower parity points were much wider than in later periods, providing scope for substantial exchange-rate variations. The Salamancan economists were quite aware of non-PPP influences on the exchange rate as determinants of the spread between parity points, citing such matters as differences among coins in metallic weight or fineness and costs of transporting coin or bullion. As was suggested above, the depreciation of Spanish currency against foreign exchange in the absence of (or correcting for) changes in these non-PPP influences provided an impetus for Salamancan development of the PPP theory.

1.1.2 Tudor Period

Gerrard de Malynes, writing in England at the end of the Tudor period, in 1601, presented a PPP theory of foreign exchange not unlike that of Navarro. The Salamancan had published his treatise 45 years earlier, but Malynes apparently was unaware of any Spanish predecessors. Two modern authors, Schumpeter (1954) and Kalamotousakis (1978), trace the origins of the PPP theory to Malynes.

Like Navarro, Malynes has all the ingredients of the PPP approach and leaves it to the reader to put them together. He begins with the quantity theory of money[6]:

> plentie of money maketh generally things dear, and scarcitie of money maketh likewise generally things good cheape.... According to the plentie or scarcitie of the monie then, generally things become dearer or good cheape, whereunto the great store or abundance of monie and bullion, which of late years is come from the west Indies into Christendom hath made euery thing dearer according to the increase of monie. (1601, 1924, p. 387)

This clear exposition of the quantity theory is at variance with the comments of Angell (1926, p. 13) that Malynes "has no clear idea of the quantity theory of money" and that "his lack of any form of the quantity theory led him into numerous errors." Even Schumpeter is unduly restrained in his acknowledgement of Malynes' accomplishment: "Malynes... tried, I think, to convey the genuine quantity-theory idea—though in a quite rudimentary form" (1954, p. 314).

Malynes then presents the money-supply theory of foreign exchange:

> plentie of money beyond the seas maketh the price of the exchange to rise, and scarcitie of money likewise beyond the seas maketh the price to fall: and so on the contrary with vs here in England, plenty of money maketh the price to fall, and scarcity of money maketh the price to rise. (1601; 1924, p. 397)

Malynes goes on to state the obvious but rarely expressed pedagogical point that, for this rule, "the head of the exchange resteth with vs," where "the head of the exchange is taken to bee at such a place or places where the price doth not alter" (1601; 1924, pp. 390–391). In other words, the exchange rate is defined as the number of units of foreign currency per English currency. The inverse definition, he notes, would reverse the direction of price movements in the theorem.

The quantity theory of money and the money-supply theory of foreign exchange together imply the PPP theory. The question arises whether Malynes was at all aware of this connection.

During the Tudor reign in England, exchange controls of various degrees of severity were periodically adopted and then removed.[7] It was paid of Malynes' genius as a mercantilist that, though he recommended officially fixed exchange rates supported by exchange control, he was concerned with the proper *level* at which an exchange rate should be fixed. He asserts that "the exchange for all places ought to be kept at a certaintie in price, according to value for value" (1601; 1924, p. 397).

It would be pleasing for the PPP theory if Malynes simply meant PPP for "value for value." However, what he seems to mean, rather, is the true *mint* parity. Yet the PPP theory may be deemed reached by another route; for Malynes has in mind a theory of price-level changes in response to exchange rates differing from the mint parity. As Schumpeter states: "he [Malynes] nicely explains how, if a country's currency falls below its mint par and coin flows out in consequence, then prices will fall in that country and rise abroad" (1954, p. 345). It is reasonable to project that the level at which Malynes recommended that the exchange rate be fixed was not only the mint parity but also (ultimately if not initially) the purchasing power parity, since specie flows and price-level changes (the price specie-flow mechanism) at home and abroad would make the PPP equal to the mint parity. This interpretation of Malynes' theory is certainly that of Schumpeter:

> When countries are in monetary equilibrium with reference to one another, then . . . gold is distributed between them in such a way that there is no profit in transferring any part of a country's holdings to any other country. We may express this by saying that the purchasing power of gold is internationally at par and also, from the standpoint of the inflation theory of foreign exchange, that this parity and its variations are the (immediately) determining factors in the foreign-exchange market. This Purchasing-Power Parity theory, or some rudimentary form of it, goes far back and can … certainly be attributed to Malynes. (1954, p. 737)

By the "inflation theory of foreign exchange," Schumpeter means precisely the PPP theory; for he writes: "We may label as Relative Inflation the variations in the value of a country's monetary unit, in relation to the value of other countries' monetary units, and speak accordingly of an Inflation Theory of Foreign Exchange" (1954, p. 736).

Malynes can thus be interpreted as seeing a role for the PPP theory whether the exchange rate is floating or fixed. Under a floating rate, PPP determines the exchange rate via the quantity theory of money and the money-supply theory of foreign exchange. Under a fixed rate, that is, one confined within specie points, the price specie-flow mechanism operates to change countries' price levels until countries' relative price levels (absolute PPP) equal the mint parity.[8]

Malynes does not draw the conclusions himself for either proposition, probably because of his overconcern with defects in the international payments mechanism, to the neglect of completing his basic arguments. For example, Wu (1939) points out that Malynes sometimes seems to assume price inelasticity of demand for England's exports, while Schumpeter (1954) sees Malynes as observing (unhappily) an unfavorable terms of trade for his home country.

Notes

1. Yet recognition of the Spanish accomplishment remains sparse in the literature. All else that I could find are one-sentence acknowledgements by Myhrman (1976), Isard (1978), and Officer (1976a). The only subsequent analysis of the Salamancan contribution is provided by Grice-Hutchinson (1978) herself in a study of Spanish economic thought over a much longer time period.

   Grice-Hutchinson notes that the School of Salamanca had been discovered earlier in the non-English literature, notably by J. Larraz (Spanish), writing in 1943. She mentions as his predecessors A. E. Sayous (French) in 1928 and Alberto Ullastres Calvo (Spanish) in 1942.
2. "Scholastic writers noticed the effect of the scarce or plentiful money supplies on exchange rates…Outstanding among them was Pegolotti's book, written about 1340, and Uzzano's book, written about a century later. Both of them were aware of the influence of the monetary scarcity (strettezza) or ease (larghezza) on exchange rates" (Einzig, 1970, p. 94).
3. "[A] statement of the PPP theory would have to involve going beyond the money-supply influence on exchange rates to that of the price level. A quantity theory of money, even in rudimentary form, would be required, which these and other writers of the time failed to have" (Officer, 1982, p. 29).
4. Historians of economic thought generally attribute origination of the quantity theory to Jean Bodin, who published his work in 1568. Schumpeter (1954) is apparently unaware of de Navarro, though he refers to later Salamancan writers on the topic.
5. All quotations are from Grice-Hutchinson's translation (1952, pp. 91–94).
6. Quotations are from Malynes' *A Treatise of the Canker of England's Commonwealth* (1601), as excerpted in Tawney and Power (1924).
7. A history of these exchange controls is provided by Einzig (1970, ch. 14).
8. With a spread between upper and lower parity points, there is no tendency for the exchange to settle mid-way, at the mint parity itself. While Malynes did not state this point, the Salamancan writers understood it; for them, PPP is the exchange-rate determinant within the spread.

   [There are two extensions to the first sentence of this note. First, with asymmetrical gold points (for example, due to differential costs of import and export gold arbitrage—see Officer, 1996, pp. 179–180), mint parity and the spread midpoint differ (see Sect. 23.2.3). Second, under ideal assumptions, "the critical exchange rate at which there is zero speculative demand and supply of foreign exchange is the midpoint of the spread, not the mint parity" (Officer, 1996, p. 197). However, the critical exchange rate need not be the PPP. In general, this critical exchange rate, the spread midpoint, and mint parity differ from each other and from the PPP exchange rate. For normative use of the critical exchange rate, see Sects. 21.2 and 23.2.3.]

References

Angell, James W. (1926). *The Theory of International Prices*. Cambridge, MA: Harvard University Press.

Einzig, Paul (1970). *The History of Foreign Exchange*. London: Macmillan, 2nd ed.

Grice-Hutchinson, Marjorie (1952). *The School of Salamanca*. Oxford: Clarendon Press.

———. (1978). *Early Economic Thought in Spain 1177–1740*. London: George Allen & Unwin.

Isard, Peter (1978). *Exchange-Rate Determination: A Survey of Popular Views and Recent Models*. International Finance Section, Princeton University.

Kalamotousakis, George J. (1978). "Exchange Rates and Prices: The Historical Evidence." *Journal of International Economics* 8 (May), pp. 163–167.

Malynes, Gerrard D. (1601). *A Treatise of the Canker of England's Commonwealth*. London: R. Field for W. Iohnes. Excerpts reprinted in R. H. Tawney and Eileen Power, *Tudor Economic Documents*, vol. 3, pp. 386–404. London: Longmans, Green, 1924.

Myhrman, Johan (1976). "Experiences of Flexible Exchange Rates in Earlier Periods: Theories, Evidence and a New View." *Scandinavian Journal of Economics* 78 (No. 2), pp. 169–196.

Officer, Lawrence H. (1976). "The Purchasing-Power-Parity Theory of Exchange Rates: A Review Article." International Monetary Fund *Staff Papers* 23 (March), pp. 1–60.

———. (1982). *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance*. Greenwich, CT: JAI.

———. (1996). *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*. Cambridge: Cambridge University Press.

Schumpeter, Joseph A. (1954). *History of Economic Analysis*. New York: Oxford University Press.

Wu, Chi-Yuen. *An Outline of International Price Theories*. London: George Routledge.

1.2 Malynes, Gerard De (fl. 1586–1623)

Originally published in *The New Palgrave Dictionary of Economics*, third edition. Palgrave Macmillan, 2018, pp. 8155–8156.

A merchant of English parentage, born in Antwerp at an unknown date, Malynes was a commissioner of trade in the Low Countries about 1586. He came to London and was frequently consulted on commercial questions by the Privy Council in the reigns of Elizabeth I and James I. He became an assay master at the mint and obtained a patent to supply farthings; he was imprisoned for a time, complaining later that he had been ruined by being paid in his own coins! He also served as a spy for England. Called on by the standing commission on trade for evidence on the state of the coinage, he published a series of pamphlets on money and prices. A mercantilist and a bullionist, he was heavily influenced by Scholastic literature.

Malynes viewed individual commodity prices as determined by demand and supply. However, he was more interested in the price level, governed by the quantity of money (Malynes 1601a, 1603). An expanding money supply, associated with a rising price level, decreased interest rates and stimulated the economy (1601a, 1622a). Therefore Malynes viewed usury as at best a necessary evil (see Muchmore 1969, p. 346) and, above all, opposed any export of specie whatsoever.

Rejecting the balance of trade theory, Malynes charged that 'bankers' (exchange dealers) controlled the exchange rate (1601a, 1622a, b, 1623). By their incorporation of usury in the price of a bill of exchange and through speculation, they conspired to undervalue sterling, leading to a deterioration in England's terms-of-trade ('overbalancing') and a specie outflow (1601a, 1622a, 1623). But overvalued sterling would not lead to a specie inflow, because the export proceeds would be spent on luxury imports (1601a). Yet Malynes (1601a) has a theory of price level changes in response to exchange rates differing from mint parity and money flowing between countries—a price specie-flow mechanism, marred only by the assumption of inelastic demand. His solution to the twin problems of specie outflow and terms of trade deterioration is comprehensive exchange control with enforced exchange dealings at rates fixed at mint parities (Malynes 1601a, 1622a, b; Muchmore 1969, pp. 347–348).

Selected Works of Malynes

1601a. *A treatise of the canker of England's commonwealth*. London: Richard Field for William Iohnes. Reprinted in part in *Tudor economic docum*ents, vol. 3, ed. R.H. Tawney and E. Power. London: Longmans, Green, 1924.
1601b. *Saint George for England, allegorically described*. London: Richard Field for William Tymme.
1603. *England's view, in the unmasking of two paradoxes*. London: Richard Field.
1622a. *Consuetudo, vel lex mercatoria, or the Ancient Law-merchant*. London: Adam Islip.
1622b. *The maintenance of free trade*. London: I. Legatt for W. Sheffard.
1623. *The centre of the circle of commerce*. London: William Iones.

References

Johnson, E. 1933. Gerard de Malynes and the theory of the foreign exchanges. *American Economic Review* 22: 441–455.
Muchmore, L. 1969. Gerrard de Malynes and mercantile economics. *History of Political Economy* 1: 336–358.
Officer, L.H. 1982. The purchasing-power-parity theory of Gerrard de Malynes. *History of Political Economy* 14: 251–255.
de Roover, R. 1974. *Business, banking and economic thought*. Chicago: University of Chicago Press.
Schumpeter, J. 1954. *History of economic analysis*. Oxford: Oxford University Press.
Spiegel, H. 1971. *The growth of economic thought*. Englewood Cliffs: Prentice-Hall.
Wu, C.-Y. 1939. *An outline of international price theories*. London: George Routledge & Sons.

CHAPTER 2

Gustav Cassel

Originally published in *Purchasing Power Parity and Exchange Rates*, JAI, 1982, pp. 85–102.

2.1 From Salamancans/Malynes to Cassel

Purchasing power and Gustav Cassel: the names are inextricably linked. Yet, between the Salamancans/Malynes (in the sixteenth century and 1601) and Cassel (in World War I), a considerable number of authors postulated the PPP theory. A summary of their work follows.

In 1761, during the Swedish bullionist period, Pehr Niclas Christiernin made explicit the concept of the price level—crucial for PPP. This distinguishes him from all other pre-twentieth-century proponents of the PPP theory, who, of course, had to understand the price-level concept in order to present the PPP theory, but did not formulate the concept in as explicit terms. In contrast to the views of other interpreters, I argue that Christiernin had both an absolute and a relative PPP theory (Officer 1982, pp. 37–42). The sole deficiency is that the commodity prices considered pertain only to the domestic country (Sweden). In this respect, Christiernin's analysis is inferior to that of Malynes, who incorporates price changes both at home and abroad.

Jean-Baptiste Mosneron, during the French Revolutionary period, stated a relative-PPP theory, with the same deficiency as that of Christiernin (Officer 1982, pp. 42–43).

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_2

Henry Thornton was an English bullionist who, writing in 1802, presents a PPP theory that again ignores the foreign price level. He goes beyond all previous PPP writers in recognizing that the increase in domestic prices (which causes a depreciation of the country's currency in the foreign-exchange and bullion markets) need not arise solely from an expansion of the money supply. Any other cause of an increase in the domestic price level also leads to the PPP result of currency depreciation (Officer 1982, pp. 47–51).

Francis Horner, in reviewing Thornton's work, exposits a clear statement of the relative-PPP theory and is the first writer to distinguish between, or at least use the terms, "real" and "nominal" exchange rates (Officer 1982, pp. 51–53). With Thornton and Horner primary authors of the 1810 Parliamentary Bullionist Report, it is not surprising that this famous document offers a precise statement of the relative-PPP theory (Officer 1982, pp. 69–71).

A third English bullionist, John Wheatley, provides the first complete formulation of the PPP theory. He had in mind a firmer concept of the price level than predecessors, as he demonstrates understanding of an *index-number* representation of the price level. He presents a complete two-country formulation of the PPP theory. In operating without the assumption of a constant foreign price level, Wheatley goes beyond his immediate predecessors and reverts to the less-restrictive, two-country formulation of the Salamancans and Malynes (Officer 1982, pp. 53–61).

John Leslie Foster, an Irish bullionist, asserted the PPP theory in 1804. The 1810 work of William Blake, another bullionist writer, centers on the distinction between real and nominal exchange rates. He anticipates modern PPP analysis in postulating that, in the short run, both PPP and real factors determine the market exchange rate, whereas, for the long run, Blake is a true believer in PPP (Officer 1982, pp. 61–64).

David Ricardo made many original contributions to economics, but his treatment of PPP, in 1810 and 1811, went little beyond that already reached by his contemporaries. There is no doubt, however, that Ricardo was a firm believer in PPP. In view of Ricardo's fame, he was perhaps the most conspicuous proponent of the PPP theory prior to the twentieth century. Cassel (1922) credits Ricardo with the first "scientific" theory of the foreign-exchange market, and recognizes his anticipation of Cassel's PPP theory (Officer 1982, pp. 64–69). Cassel was thereby unfair not only to the other English bullionists (Thornton, Horner, Wheatley, Foster), who had prior claim on the theory, but also to earlier writers in Spain (the Salamancans) and England (Malynes), and to Swedish and French bullionist authors (Christiernin and Mosneron).

Following the tumultuous PPP literature of the English bullionist period, nearly a full century was to pass before PPP analysis was again in the forefront of theories of exchange-rate determination. In the meantime, the PPP approach was confined to four economists: John Stuart Mill, Viscount Goschen, Alfred Marshall, and Ludwig von Mises.

Mill's exposition of the PPP theory (in 1848) is less advanced than some earlier literature. He presents only the relative—and not the absolute—PPP theory. He considers a change in the price level in only the domestic, and not also the foreign, country. The real-versus-nominal exchange-rate distinction is implicitly assumed, rather than developed or extended (Officer 1982, pp. 73–76).

Writing in the 1860s, Viscount Goschen presents the PPP theory for an inconvertible paper currency (floating exchange rate) through the usual mechanism: the quantity theory of money leading to a relative-PPP result (Officer 1982, pp. 76–77).

In testimony before Royal Commissions in the late nineteenth century, Alfred Marshall displayed both the PPP theory, in absolute form, for a floating exchange rate, and its weaker variant, the law of one price, under a fixed exchange rate. More than a quarter-century later, in 1923, Marshall saw fit to reprint verbatim that part of his testimonies relating to PPP. It is intriguing that he does not take the opportunity to mention the intervening (and great!) work of Gustav Cassel. Marshall was toward the end of a long life and career. Perhaps he had not kept up with the literature. Possibly the issue of PPP was deemed too obvious and/or too unimportant to warrant more than reprinted testimony in an appendix in a new book. The unfairest speculation is that Marshall was reciprocating Cassel's neglect to mention Marshall's antecedent work on PPP.

Marshall's original contributions were imposition of balance-of-trade equilibrium as a condition for the absolute-PPP theory to hold, inclusion of international capital flows as an inhibitor of a strict-PPP determination of the exchange rate, and a careful analysis of adjustment to the law of one price under a rate constrained between specie points (a "fixed" exchange rate). All discussed in Officer (1982, pp. 77–79).

Prior to World War I (in 1912), Ludwig von Mises, writing in German, has only the relative—and not the absolute—PPP theory, with his contribution the proposition that exchange-market adjustment to an altered PPP can be either immediate or delayed (Officer 1982, pp. 80–81).

Except for the four cited authors, the PPP literature was basically dormant from the end of the Bank Restriction Period to World War I. The reasons are several: the role of Goschen in the 1860s in shifting attention from the PPP theory to the balance-of-payments theory of exchange-rate determination, the dominance of the gold standard by 1880, the narrowing of specie points during the century, the emergence of speculation to explain variations in floating exchange rates (Officer 1982, pp. 81–84). By the time of World War I, mainstream economics literature had ceased to incorporate the theory. Not for the first time in its history, the PPP theory awaited rediscovery.

Gustav Cassel rediscovered the PPP theory after the theory's long-dormant period. That is the proximate reason why it is Cassel, writing during World War I, whose name is almost invariably the first connected to the theory. Indeed, the PPP theory is sometimes called, simply, "Cassel's theory."

2.2 The Impact of Cassel

> If these critics, who express themselves in such vague general terms, were allowed to have their own way, the entire theory of the purchasing-power parity would have to be thrown to the winds, and we should be left in as much doubt as ever as to the real basis of the rates of exchange. (Cassel 1924, p. 68)

Einzig (1970, p. 264) states the reason for the association of Cassel's name with PPP:

> Neither Ricardo nor any of the earlier economists had succeeded in developing the purchasing power parity theory sufficiently, or in making a strong enough impression with their exposition of that theory, to ensure its adoption by textbook-writers before the first World War. Cassel succeeded in doing so to a remarkable degree.

One should go further than Einzig's statement. The impact of Cassel on bringing about the economics-profession's awareness of PPP analysis was greater than that of any other person in the history of PPP development. No predecessor of Cassel, no contemporary, no later individual implanted the PPP theory so broadly and firmly in the economics-profession's domain and also the public domain. Cassel was the supreme publicist of the PPP approach. Indeed, it is fair to describe that successful promotion effort as perhaps his most important contribution to the PPP literature. After Cassel, the PPP approach was sometimes neglected, but never forgotten.

What were the reasons for the success of Cassel's publicity work on behalf of PPP? First of all, no other author—again, not before Cassel, contemporaneously, or after him—wrote so prolifically on the topic. Cassel devoted a total of at least 25 English-language publications, in whole or in part, to the PPP approach. This list includes 15 articles: seven in the *Economic Journal* (1916a, b, 1917, 1918, 1919, 1920a, 1928c); five in Skandinaviska Kreditaktiebolaget *Quarterly Report* (1923a, 1924, 1925b, c, d); and one each in *Annals of the American Academy* (1920b), *Economica* (1923b), and the *Encyclopedia Britannica* (1926). Also included in these writings are eight books authored by Cassel (1916c, 1921, 1922, 1925a, 1928a, 1932a, 1932b, 1936) and two volumes in which he served as a contributor (1925e, 1928b).

The second reason for Cassel's success as a PPP publicist is that he disseminated all his PPP writings in the English-language mainstream of economics. His works on PPP were either published directly in English or translated into that language.

Third, Cassel exposited the PPP approach in an extremely forceful and assertive style of writing. It is apparent to the reader that Cassel is both exuberant about the explanatory power of the PPP hypothesis and determined to carry out a mission of replacing other exchange-rate theories with the PPP theory in the professional and public domains. This tone of Cassel's writings on PPP contrasts with the matter-of-fact or coldly analytical treatment of his immediate predecessors, that is, Mill, Goschen, Marshall, and Mises. Even the English bullionist economists did not exhibit anything like Cassel's verve, excitement, and sense of mission in their PPP writings.

Fourth, Cassel's wartime and postwar publications on the subject stimulated a controversy on the merits of PPP analysis that was even more powerful than the bullionist controversies of earlier centuries. Unlike the several bullionist debates, the reaction to Cassel in the economics literature was international in scope—with publications not only in England (the place of Cassel's earliest writings on PPP), but also in several European countries and the United States. This controversy also differed from the previous debates in its exclusive focus on PPP analysis.

Fifth, the sheer number of publications involved in the controversy exceeded, by my count, the world total of all previous writings on PPP. It is not surprising that, in Officer (1982), following essentially the present chapter—devoted to Cassel's contribution to the development of PPP theory—I abandon the approach of considering, chronologically, each and every publication dealing with PPP.

Sixth, Cassel wrote at a time of *international* ferment in foreign-exchange markets. The Swedish, French, and English/Irish bullionist experiences, in contrast, were purely domestically oriented events. The background to Cassel's publications was the First World War and the early postwar period, with inflation and large-scale exchange-rate movements spanning a large number of economies.

Seventh, Cassel gave the PPP theory its name. He did not do this until his fourth article (1918) on the topic, where the term "purchasing power parity" was used for the first time. In his first article (1916a), the equivalent term "theoretical rate of exchange" was employed. As a descriptive device, "purchasing power parity" explains the theory precisely and it immediately was adopted by the economics profession. The nominal-versus-real exchange-rate *terminology* invented by Horner and Blake a century earlier was entirely supplanted and, in fact, had never been fully incorporated in the literature.

Finally, the substantive contributions of Cassel to PPP analysis cannot be overlooked as a powerful reason for his successful implantation of the theory firmly in the economics consciousness. There were predecessors to Cassel in developing the PPP approach, but he was the first to place PPP within so systematic a framework that a clearly operational theory resulted. He distinguished carefully between the absolute and relative versions of the theory, although he did not provide them with names.[1] Also, Cassel was the first to express the theory formally in terms of statistical averages of prices. Not only did Cassel make PPP an operational theory, but also he was the first to use PPP to obtain estimates of exchange-rate disequilibria and the first to test the theory empirically. In this chapter, only theoretical aspects of Cassel's work on PPP are considered; his empirical use of PPP and his testing of the theory are discussed in Officer 1982, Part III. Further specific contributions of Cassel to PPP theory are presented in later sections of this chapter.

Persuasive as Cassel was, his PPP work gave rise to criticism as well as adoption. For surveys of the decade or more of published conflict between critics of PPP analysis (opponents of Cassel) and its proponents (followers of Cassel), the studies of Angell (1926), Ellis (1934), and Einzig (1970) may be consulted.[2]

2.3 Motivation for Cassel's PPP Theory

As far as Cassel was concerned, the *proximate* motivation for his theory was the dislocations of exchange rates during the World War combined with his disagreement with the general view that prewar exchange values of currencies (mint parities under the gold standard) would be reestablished after the war. In one of his earliest works, he writes: "All the combatants wish the world to believe that after the war their currencies will resume their normal value. But, in all probability, this problem will possess far different features from those it now presents" (Cassel 1916c, p. 57).

This situation occurred, according to Cassel, in an atmosphere of ignorance concerning the determination of exchange rates. "The discussion on the variations in exchange rates and their true explanation, which has been going on the whole world over since the outbreak of the War, has been chiefly characterized by a remarkable lack of clearness on the question as to what really determines the exchange rate between two independent currencies" (Cassel 1922, pp. 137–138). Twelve years after his initial publication on PPP, Cassel (1928a, p. 24) reflected on the background for his theory:

> During the War it was generally believed, and even officially preached, that exchanges were only disturbed by the obstacles which the War put in the way of international trade, and that exchanges must therefore be expected to revert to their normal pre-War levels as soon as peace was in sight. The fundamental wrongness of this view was made manifest by the Purchasing Power Parity theory.

There is evidence, though, that Cassel had developed the PPP theory even before the war. Meinich (1968, p. 159) declares: "Cassel says that he got the principal ideas of this theory [PPP) during his lectures at Stockholms Högskola in 1905." The reference is to a footnote in the Swedish (but not the English) edition of Cassel's *Theory of Social Economy.* Also, in beginning his first article on PPP, Cassel confirms that he possessed the theory in the prewar period; for he presents "the theory of the foreign exchanges *which I have given for some years in my lectures*" (Cassel 1916a, p. 62; italics added). If Cassel's lecture notes, perhaps as taken by a student, were found, they might indicate the precise timing of his origination of the PPP theory.

Was Cassel influenced by his antecedents in the PPP literature? In only one of his 25 publications on PPP does he allude to predecessors; only Ricardo, Mill, and Goschen are acknowledged. It is strange that Cassel does not mention the PPP discovery during the Swedish bullionist controversy of the eighteenth century; for his fellow Swede, Sven Brisman, noted that fact during Cassel's lifetime.[3]

Cassel (1922, p. 170) credits Ricardo with "the first theory of exchanges of a scientific character." He acknowledges that Ricardo applied the relative-PPP theory to a floating-exchange-rate situation and also sees him as having the law of one price (subject to transport costs) under a fixed-rate, metallic system. His reference to Goschen indicates that Cassel sees him as having a PPP-like theory only for a metallic standard, with the exchange rate under a paper standard determined by the demand for and supply of foreign exchange. This interpretation is incorrect; Goschen did apply the PPP theory to a paper currency.[4]

As for Mill, Cassel does not appear to credit him with any PPP theory at all. Mill is correctly viewed as having the "nominal" exchange rate determined by the amount of currency depreciation. However, quite unfairly, the latter is interpreted by Cassel as referring to depreciation with respect to the metal parity and not (or not also) the rise in the price level.

So Ricardo is considered by Cassel to be his true predecessor. Cassel (1922, p. 172) writes: "Ricardo finally draws various conclusions which in reality contain much of what a true theory of exchanges should contain." Perhaps other contemporary bullionist proponents of PPP, such as Wheatley and Blake, were ignored by Cassel because Ricardo was the most-prominent English economist of that time.

2.4 Quantity Theory of Money

According to Schumpeter (1954, p. 737), Cassel's PPP theory (as that of Ricardo) appeared "in characteristic association with a strict (and crude) quantity theory." It is true that Cassel adhered to the quantity theory of money throughout his writing. In fact, he was regarded as a quantity theorist of Irving Fisher's stature! However, Cassel never expounded a *simple* quantity theory. Even his first article on PPP qualifies the theory: "Now, according to the quantitative theory of money the general level of prices varies, other things being equal, in direct proportion to the quantity of the circulating medium in a country" (Cassel 1916a, p. 62).

In his subsequent writings, Cassel explained what "other things" are kept equal. Changes in output affect the price level (in the opposite direction), but money-supply changes are by far the more-important factor. Changes in the demand for money (sometimes taking the form of changes in the velocity of circulation, in Cassel's later work) are explicitly considered, including effects on the price level as well as on real output. Also, causation can move in the opposite direction, with velocity affected by changes in the price level or the money supply.

Even the essence of the quantity theory admits of a two-way causation, as the demand for money increases proportionately to the rise in prices and the supply of money passively adjusts to the demand. However, Cassel is emphatic that the initiating cause of any inflationary process is always increases in the money supply, what he calls "the creation of artificial purchasing power." Continually increasing the money supply results in *inflation*, as distinct from a once-and-for-all change in the price level. It is always within the power of the government to restrict the money supply and thereby stabilize the price level.

Cassel states the unqualified quantity-theory relationship only as a long-run proposition: "In the long run, of course, the internal purchasing power of a currency must, after all, always be determined by the amount of money in circulation" (Cassel 1925d, p. 56).

In summary, far from being a believer in a crude quantity theory, Cassel was a *sophisticated monetarist* ahead of his time. As Holmes (1967, p. 688) writes: "The idea that changes in the monetary sector would cause changes in the non-monetary sector is expressed so often in Cassel's writing…that it is amazing that one could think of him as a naive theorist—quantity or otherwise."

2.5 Price-Level Concept

Cassel is well aware of the index-number problems involved both in computing a domestic price index and in constructing relative price levels (absolute PPP). Relative-price changes and movements in the general price level, he notes, are commingled. A price index seeks to measure "how far a shifting of the center of gravity of the price-level has taken place" (Cassel 1932b, p. 463). Cassel notes that, in generating a relative-PPP measure, the index-number problem is compounded.

In constructing absolute PPP, one can obtain a precise comparison of price levels in the two countries only in the limiting case of all individual prices in one country differing from corresponding prices in the other by the same multiplicative factor; that is, all relative prices are the same in the two countries. In that case, Cassel correctly points out, there will be no international trade. With differing relative prices in the countries (the realistic circumstance), only "an approximate comparison between the purchasing power of the one currency and the other" is possible (Cassel 1928b, p. 8).

What price measure should be used in the PPP theory and for PPP computations? The issue is discussed mainly in terms of price indexes (with reference to the quantity theory of money and relative PPP), rather than in terms of price levels (absolute PPP). In forming a price index, price relatives of individual commodities are logically weighted according to their importance. Cassel suggests that the weighting pattern reflect either production or consumption of commodities.

In particular, price indexes limited to traded goods (exports and imports) are emphatically rejected. Several reasons for this decision are provided. First, such indexes "are limited to a small class of commodities, and are therefore subject to variations" (Cassel 1922, p. 47) that presumably would not be present in a broader-based price measure. Second, traded and nontraded goods are not unvarying collections of commodities. "There is never a definite group of commodities that can be exported. Even a small alteration in the rate of exchange may widen or restrict the group of exportable goods" (Cassel 1928a, p. 33).

Third, Cassel hints that the "law of one price" applies to traded-goods prices, so that they tend to move together in different countries irrespective of the amount of deviation of exchange rates from their purchasing power parities. He writes: "if export commodities have risen in relative value in the exporting country, they have probably in the importing country also risen in desirability, and therefore in value, as compared with other commodities. The higher price of the export commodities, therefore, need not necessarily cause the value of the exporting country's exchange to be reduced on a like scale" (Cassel 1922, p. 155).

Cassel contends that a *general* price level is required to define absolute PPP and a *general* price index is needed for relative PPP. He writes: "the height of the general price level in different countries…[is needed] to make a real calculation of the purchasing power parities" (Cassel 1922, p. 182) and "The whole theory of Purchasing Power Parity essentially refers to the internal value of the currencies concerned, and variations in this value can be measured only by general index figures representing as far as possible the whole mass of commodities marketed in the country" (Cassel 1928a, p. 33).

The most-logical interpretation of the general price index envisaged by Cassel for relative PPP would be the gross-domestic-product (GDP) deflator and its analogue—the GDP price level—for absolute PPP. Certainly, the very concept of the internal purchasing power of a country's currency implies that Cassel means to exclude import prices from the measure and to include export prices. National accounts had not yet been developed at Cassel's time of writing, so the precise concept of a GDP price measure was alien to him. He declares that one must consider all available general price indexes: the wholesale price index, retail price index, cost-of-living index, and wage-rate index.

At one point, Cassel suggests that the wholesale price index is most suitable to measure long-run price movements, the reason being: "We must confine ourselves to typical standard commodities of a practically fixed quality" (Cassel 1932b, p. 463). This passage is clearly an aberration. Elsewhere, Cassel notes: "An index for wholesale prices may be based on statistics of general prices or else on statistics of the prices of import and export goods" (Cassel 1922, p. 47). The latter-type index is rejected.

Ultimately, Cassel leans in favor of a cost-parity concept. He writes: "Only when prices have adjusted themselves to one another so as to make prices of products correspond to their cost of production, can we regard the usual index number of wholesale prices as a fairly reliable index of the movements of the general level of prices" (Cassel 1921, p. 110), This statement is supported elsewhere: "The level of wages in the country, therefore, is always a very important factor—in the long run may be the predominating one—in determining the international value of the country's currency" (Cassel 1922, p. 144).

2.6 Absolute-PPP Theory

Cassel's theory of PPP is appropriately named; for its foundation is the idea that the value of a currency—and therefore the demand for it—is determined fundamentally by the amount of goods and services that a unit of the currency can buy in the country of issue, that is, by its "internal purchasing power." The internal purchasing power of a currency is sometimes called simply its "purchasing power" or—as in Cassel's early writings—its "buying power" or "paying power." Irrespective of the term used, the domestic purchasing power of a country's currency is defined as the inverse of the price level.

One of Cassel's many contributions to PPP analysis is that he was the first author to formulate his theory in terms of a schemata much-later outlined in Officer (1982, chapter 2). The long-run equilibrium exchange rate— called by Cassel the "equilibrium rate of exchange," "normal rate of exchange," "equilibrium position" of the exchange rate, "normal position" of the rate, "normal parity"—is defined as the value of the exchange rate that yields balance-of-trade equilibrium. It is fair to interpret Cassel's balance-of-payments concept more broadly so that he means equilibrium in the current account. He writes: "The main reason why we pay anything for a foreign currency is of course that this currency represents in the foreign country a purchasing power which can be used for acquiring the goods or for paying for the services of that country" (Cassel 1926, p. 1086). For simplicity, Cassel sometimes assumes that trade consists entirely of "commodities," that is, goods rather than services. This procedure justifies use of the trade rather than current account as the balance-of-payments concept in the definition of the equilibrium exchange rate.

Cassel points out that if the actual exchange rate (price of B-currency in terms of A-currency) exceeds (falls below) the equilibrium rate, then country A would have a trade surplus (deficit). Only when the actual rate is at equilibrium is there a trade balance. Now, the principal—though not the sole—determinant of the equilibrium exchange rate is the ratio of the internal purchasing powers or price levels of the two countries, that is, the absolute PPP. Though Cassel does not use the adjective "absolute," he defines and uses the absolute-PPP concept correctly and consistently.

Why is the PPP the main determinant of the equilibrium exchange rate? Since the value of a given currency is basically determined by its domestic purchasing power (inverse of the price level), the equilibrium value of one currency (relative to another) is fundamentally determined by, and in a limiting circumstance equal to, the ratio of the internal purchasing powers of the currencies, that is, the (absolute) PPP. As Cassel (1928b, p. 7) writes: "Obviously, in the state of equilibrium a certain sum of money must have about the same purchasing power if converted into the one currency or into the other."

In Cassel's most-thorough analysis, there is both an f function and a g function, in the terminology of Officer (1982, chapter 2).[5] The long-run equilibrium exchange rate tends to equal the PPP (the f function)—"The internal purchasing power of the two currencies contemplated determines only the equilibrium of the rate of exchange" (Cassel 1926, p. 1086)—and the actual exchange rate (short-run equilibrium exchange rate) under a paper standard and floating exchange rate tends to equal the long-run equilibrium rate (the g function). In each case (f function, g function), there may be deviations of the dependent variable (long-run equilibrium rate, actual rate) from its ultimate determinant (PPP, long-run equilibrium rate). These deviations are discussed in Sect. 2.9.

As a short-cut in his analysis, Cassel *equates* the PPP with the equilibrium exchange rate. "In order to emphasize this dominating influence of the internal purchasing power in fixing the equilibrium rate of exchange, we call this rate, as here defined, the purchasing power parity between the two currencies" (Cassel 1926, p. 1086). He combines this simple f function with the g function to obtain an h function in which the exchange value of a floating currency is a function principally of the PPP and, in the limiting case, equals the PPP. As Cassel (1916a, p. 62; 1932b, p. 513) writes:

> Thus the rate of exchange between the two countries will be determined by the quotient between the general level of prices in the two countries...Thus the price of the bill on country B must, as an expression of the value of the currency of country B in terms of the currency of country A, be directly determined by the relation existing between the value of money in countries B and A respectively. This relation is the purchasing power parity of the two currencies.

Cassel states a neutrality theorem for the absolute-PPP theory. High prices within a country will not encourage imports or discourage exports, as these prices will be counterbalanced by a low exchange value of the domestic-country's currency, and the equilibrium balance of trade is maintained. Similarly, the level of the exchange rate is irrelevant for real behavior, providing only that the exchange rate reflects the PPP. So Cassel (1920b, p. 262; 1922, p. 157) can comment: "In reality the purchasing power parity represents an indifferent equilibrium of the exchanges in the sense that it does not affect international trade either way…But as soon as this parity [PPP] has been established at a certain level it is of no importance whether this level is high or low."

An effective adjustment mechanism preserves the tendency of the exchange rate to equal the PPP. Cassel uses the term undervalued (overvalued) exchange rate to denote an exchange value of a country's currency below (above) its PPP. An undervalued (overvalued) exchange rate encourages (discourages) exports and discourages (encourages) imports, thus increasing (decreasing) the demand for the country's currency in the foreign-exchange market and restoring the equality of the exchange rate with the PPP. At this time, what Cassel calls the "artificial" stimulus or hampering of trade, ceases and equilibrium in the balance of trade is restored.

Though not formally using the elasticity concept, Cassel is an "elasticity optimist," believing in high price elasticities for exports and imports and, therefore, in a relatively large response of the balance of trade to a change in the exchange rate. His belief is especially strong for countries at a high level of development engaged in close commercial relations. Not only will a small deviation of the exchange rate from the PPP significantly affect the amount of trade in existing commodities, but also previously untraded commodities will become exported or imported (and some kinds of previously traded commodities will cease to be imported or exported).

High elasticities imply a great stability to the exchange rate at the equilibrium (PPP) level. Cassel (1932b, p. 661) declares: "the rate of exchange in its equilibrium position—always on the assumption of a constant value of money—possesses a great stability, that is, a great power of resistance against changes in the real conditions of international trade which tend to shift the rate in one direction or the other."

The ability to use currency to purchase goods and services in the country of issue is the foundation of Cassel's PPP theory. So he notes that the theory works best—that is, that the short-run equilibrium exchange rate is expected to have minimum deviation from the PPP—under conditions of free international trade. Cassel also states that the theory holds when trade restrictions have equal impact in both directions, that is, on both imports and exports of a country.

Under normal conditions, Cassel's theory involves a strict direction of causation, from a country's money supply to its price level and thence (given the foreign price level) to the exchange rate. "The sequence of cause and effect is incontestable" (Cassel 1924, p. 68). In particular, a rise in the foreign price level cannot affect the domestic price level, providing the exchange value of the domestic-country's currency appreciates in the same proportion as the PPP (ratio of the foreign to the domestic price level) rises.

However, Cassel mentions several exceptions to the strict chain of causation. First, if the domestic currency is undervalued (overvalued) on the foreign-exchange market with respect to the PPP, then imports are made more expensive (cheaper) domestically and exports encouraged (discouraged) because of their lower (higher) price in foreign currency. The higher (lower) price of import goods spreads to the general price level, and the increased (reduced) exports also acts to increase (decrease) the price level. Cassel notes that it is still within the power of the country to prevent the stimulating (depressing) effect on the domestic price level by suitably controlling the money supply, restricting or expanding it as the case may be.

Second, in a period of moderate and relatively stable inflation, the valuation of the exchange rate will anticipate the future currency depreciation over, say, the next year or several months (rational expectations). In principle, notes Cassel, the PPP theory still holds; as the exchange rate is affected by the expected domestic (relative to foreign) price level.

The third case of reverse causation occurs under hyperinflation.[6] In this situation, "the causal connection between the rise of prices and the rate of exchange is reversed, that is to say, the falls in the rate of exchange now become the basis for new rises of prices" (Cassel 1924, p. 69). The reasons are that the domestic currency becomes subject to adverse speculation by foreigners and that the currency becomes replaced by foreign currencies in its domestic roles of medium of exchange and unit of account. When the currency loses its domestic functions, Cassel observes, one cannot reasonably expect the PPP theory to be applicable.

It goes without saying that Cassel rejects the balance-of-payments approach to exchange-rate determination, in particular, "the popular fallacy that the movements of the exchanges could be explained by the balance of trade" (Cassel 1920a, p. 44). The adjustment mechanism that makes PPP a stable equilibrium value of the exchange rate would correct any undervaluation (overvaluation) engendering a balance-of-trade surplus (deficit).

Another argument Cassel employs against the balance-of-payments theory involves, in effect, expanding the concept of payments balance underlying the equilibrium exchange rate from the trade or current account to the basic (or perhaps official-settlements) balance (though these payments terms are, of course, not used). He declares that a deficit or surplus on current account would be fully compensated in the capital account. Cassel believes in a well-functioning transfer mechanism; so that a current-account deficit (surplus) is balanced by a surplus (deficit) on autonomous capital account: "For if a country buys more from another than it sells to it, the balance must be paid in some way; say, by export of securities or by loans in the other country. Thus the balance of payments must on the whole equalize itself, and there is no reason for a definite alteration in the rates of exchange" (Cassel 1921, p. 47).

In the language of the transfer problem, Cassel (1928b, pp. 17–18) states: "a real transfer of capital will not affect the equilibrium of the rate of exchange, which will continue to be determined by the Purchasing Power Parity." He writes: "an export of capital is always counterbalanced by an export of goods to the same value. Goods may, of course, be replaced by services" (Cassel 1928b, p. 20). This statement is a good indication that Cassel had a basic-balance payments concept in mind.

What if there is a fixed exchange rate (for Cassel, taking the form of the gold standard) rather than a floating rate? Purchasing power parity remains the principal determinant of the exchange rate. If PPP represents the long-run equilibrium exchange rate, then it must be contained within the gold points. Otherwise, over time, the country will either gain or lose international reserves without limit. In a passage vaguely anticipated by Malynes, Cassel (1928a, pp. 31–32) writes:

> The purchasing power of each currency has to be regulated so as to correspond to that of gold; and when this is the case, the Purchasing Power Parity will stand in the neighborhood of the gold parity of the two currencies. Only when the purchasing power of a currency is regulated in this way will it be possible to keep the exchanges of this currency in their parities with other gold currencies. If this fundamental condition is not fulfilled, no gold reserve whatever will suffice to guarantee the par exchange of the currency.

Cassel argues that what caused an exchange-rate change under a floating rate now brings about a corresponding change in the domestic price level under a fixed rate. This maintains the law of one price, though not the strict direction of causality postulated by the PPP theory for a floating rate. Ultimately, though, even under the gold standard, the country can determine its price level by controlling its money supply.

Cassel rejects the modern monetarist view that a country completely loses the ability to determine its money supply under a fixed rate. He writes:

> But it would be impossible to keep up the gold standard if the purchasing powers of the currencies were not maintained at a corresponding level and if the supply of means of payment in both countries were not regulated to that end. (Cassel 1926, p. 1086)

2.7 Relative-PPP Theory

Cassel's PPP theory is basically a theory of *absolute* PPP. He justifies a theory of relative PPP on the *empirical* grounds that measures of price levels—required to apply the absolute-PPP theory— are virtually impossible to obtain. It is much easier to use a relative-PPP approach, since the only price data required are measures of inflation (price index numbers) in the countries considered. He writes: "We have no trustworthy measure for the absolute purchasing power of a currency in its own country. With index numbers, we are only able to determine the relative changes in this purchasing power from time to time" (Cassel 1932b, p. 660).

Cassel's theory of relative, like that of absolute, PPP is consistently presented throughout his writings. A succinct statement of his theory is "the rates of exchange should accordingly be expected to deviate from their old parity in proportion to the inflation of each country" (Cassel 1918, p. 413). A comprehensive description of his theory begins with the actual exchange rate in a base period, which must be a "normal" period. This exchange rate is multiplied by the ratio of proportionate changes in price levels in the countries concerned. The result is the (relative) PPP in the current period.[7] The ideal base period for Cassel is one in which the exchange rate is at its equilibrium level, best of all when that level is the absolute PPP in the base period. He writes that one must:

> start from a given equilibrium at a time when the exchange rate is presumed to be known, and on the basis of this rate calculate that rate which corresponds to the same equilibrium if an inflation of the currencies has taken place without any change having otherwise occurred. (Cassel 1922, pp. 175–76)

The question arises as to whether the PPP so calculated, that is, the relative PPP in the current period, is equal to the absolute PPP newly calculated for this period, presumed to be the new equilibrium exchange rate. The answer is affirmative, according to Cassel, only if the changes in the economies that occurred since the base period were purely monetary in nature. Cassel notes that real changes may occur in this connection: "Strictly speaking, one must take into consideration the possibility that the normal levels [of exchange rates] might be altered somewhat as a result of changes in the entire economic situation of the countries in question, and also in the conditions of trade between them" (Cassel 1932b, p. 515).

Cassel correctly points out that real changes in an economy will be associated with changes in relative prices. Only under a uniform inflation, where all prices change proportionately, is the calculated relative PPP necessarily equal to the new absolute PPP. This is a neutrality hypothesis for relative PPP. As Cassel (1922, pp. 141–142) writes:

> If in each country prices are unaltered in their relation to one another, but have only undergone a common rise, then there is nothing to prevent our supposing the balance of trade between the countries to be unaltered. The equilibrium of the exchanges must, then, have been dislocated in the manner shown by the ratio of the deterioration of money in the two countries. If, on the other hand, the different prices have moved in their relation to one another, this circumstance may possibly in itself have affected the equilibrium of international trade and have caused some dislocation of the equilibrium of the exchanges.

2.8 Digression: Error Term in PPP Theory

It is a ridiculous caricature of PPP to formulate that the exchange rate cannot deviate even temporarily from the PPP. In a sense, this relationship is the most-extreme form of the PPP theory; but it has never been advocated by a proponent of the theory. Yet Nobel Laureate Paul A. Samuelson (1964) attacked the PPP literature on the grounds that it posits an unqualified equality between the exchange rate and the PPP. Considering the equation:

$$\text{R index} = (\text{American Export Price Index})/(\text{European Export Price Index}) \quad (2.1)$$

where R is the exchange rate, and assuming that the United States exports good 3 and Europe good 1, he comments: "Obviously, a point-of-time equality like (2.1) is complete nonsense, since $R = P_3/p_1$ is like saying that the \$2.80 price per £ must equal the ratio of the price of a California sherry to the price of a European Volkswagen" (Samuelson 1964, p. 149). More generally, Samuelson (1964, p. 153) concludes:

> Unless very sophisticated indeed, PPP is a misleadingly pretentious doctrine, promising us what is rare in economics, detailed numerical predictions. Few doubt that long-run wheat prices are determined by supply and demand equations rather like the one above [not shown here]; but whoever expects from this analysis detailed numerical predictions based upon simple historical calculations?

Fifteen years later, this position was supported by Katseli-Papaefstratiou (1979), who ends her study with the observation: "In conclusion, I am afraid there is an important element of truth in Samuelson's (1964, p. 153) statement that 'unless very sophisticated indeed, PPP is a misleadingly, pretentious doctrine, promising us what is rare in economics, detailed numerical predictions'" (Katseli-Papaefstratiou 1979, p. 29).

Samuelson's assertion that PPP theory is generally devoid of an error term is incorrect. Rather, the strictest form of the (absolute-PPP) theory postulated by proponents is the h function taking the form:

$$\text{short-run equilibrium exchange rate} = \text{PPP } \textit{plus} \text{ error term} \tag{2.2}$$

and similarly for the f and g functions, and the entirety also for relative PPP and for other functional forms (such as logarithmic).

True, advocates of strict PPP do not generally state the error term in mathematical symbols: a literary acknowledgment of a random error in the relationship might suffice. Even if a verbal discussion of an error term is absent, it is unfair to project the absurdity of an *exact* theory on PPP theorists in general, and on Cassel in particular (see Sects. 2.8–2.9). Unless a statement is made to the effect that the exchange rate equals PPP in any time period, always and everywhere, a random error term should be viewed as implicit in the relationship.[8] And this is true of Cassel's work (again see Sect. 2.9).

At the opposite extreme to Eq. (2.2), one can envisage a multivariable multi-equation explanation of the exchange rate, which includes PPP as but one variable with no overriding importance in determining the exchange rate. Decidedly that model is outside the rubric of Cassel's theory. For Cassel, PPP is the most-important—but not necessarily the sole—determinant of the equilibrium exchange rate (once more, see Sect. 2.9).

2.9 Flexibility of PPP

Cassel's form of the f, g, h functions (in the symbology presented in note 5) involves the PPP not as the only systematic variable explaining the exchange rate but rather as the most-important such variable. He allows room both for random influences and for other (though less-important) explanatory variables in the f, g, h functions. There are many ways in which Cassel makes clear that his PPP theory takes a less-restrictive form than strict equality.

First, throughout his writings, the effect of the PPP on the exchange rate is described in terms suggestive of a non-restrictive influence. Cassel states that the exchange rate is "determined essentially" or "governed essentially" by the PPP; or determined "in the main," "principally," "approximately," "in a rough sense," or "broadly speaking" by the PPP. He writes that PPP is the "essential factor" or "fundamental factor" or "dominating influence" on the exchange rate. The PPP theory is said to hold "broadly speaking" or in a "rough sense." Holmes (1967, p. 692) notes that "Cassel *always* had such qualifying phrases." While this statement is an exaggeration (note the quotations from Cassel in Sect. 2.6 on the absolute-PPP theory), it is true that it is difficult, if not impossible, to find entire passages in Cassel's work in which no qualifying language appears.

Second, there are two intriguing passages in Cassel—one relating to absolute, the other to relative PPP—in which the theory is described in weak terms indeed. He writes that absolute PPP "presents a solution of the exchange problem in only a first and quite rough approximation" (Cassel 1922, p. 139). Ten years later, he argues that relative PPP is "satisfactory for a first rough calculation of the new equilibrium level of the rates of exchange after big monetary changes have occurred" (Cassel 1932b, p. 661).

Third, Cassel allows for a random error term in the f, g, h functions, so that the exchange rate does not equal the PPP even if no other systematic influence is present. It is true that Cassel does not express his equations, and therefore their error terms, in mathematical language. Holmes (1967, p. 693) argues convincingly, however, that Cassel "did discuss random fluctuations in a literary context" and so included "randomly distributed error terms in the equations of his operational theory."

Cassel (1928a, p. 32) writes of "small fluctuations in the rate of exchange…caused by fluctuations of demand and supply of bills on the exchange market." He declares that, even if non-PPP variables that systematically influence the exchange rate are absent or dormant, there may be divergences (described as small and/or temporary) of the exchange rate from the PPP. For example, abstracting from non-PPP factors, he argues that "the rate of exchange…cannot show more than small and quite temporary deviations from this level [PPP]" (Cassel 1928b, p. 17). These are all allusions to a random-error term.

Fourth, Cassel acknowledges that there are lags in the adjustment mechanism that corrects an undervaluation or overvaluation of a country's currency with respect to the PPP. In this context, he writes: "In reality, however, this restoring of the equilibrium may take a long time, especially if the forces which keep the rate down are powerful and are continually at work" (Cassel 1922, p. 158).

Fifth, Cassel makes the general qualification that, in principle, any real change in the economy can affect the exchange rate. "Theoretically, any change in the economic conditions in the two countries or in the trade relations between them may cause an alteration in the rate of exchange" (Cassel 1928c, p. 589). He argues that real changes ("the effects of economic causes on the rate of exchange") are generally dominated by monetary changes ("those of monetary causes, i.e., of alterations of the price levels"):

> Alterations of the price level in one country may easily cause the rate of exchange to rise ten or a hundred times or even much more above its former height; whereas, if the general levels of prices in both countries remain constant, only extraordinary perturbations of the economic conditions are likely to call forth movements of the rate of exchange of any practical importance. (Cassel 1928c, p. 590)

Sixth, Cassel explicitly discusses the non-PPP variables in the f, g, h functions. He provides a large number of reasons why a floating exchange rate may systematically diverge from the PPP. These reasons may be summarized as follows[9]:

1. Trade restrictions may be more severe in one direction than in another. For example, if a country's imports are more restricted than its exports, the exchange value of the country's currency may exceed the PPP.
2. Differences in countries' situations regarding transport costs may also cause the exchange rate to diverge from the PPP.
3. It is possible that speculation in the foreign-exchange market is against a country's currency and therefore reduces the currency's exchange value below the PPP. However, speculation usually plays a stabilizing role in the exchange market, moderating fluctuations in the exchange rate.
4. Anticipated future inflation in a country may lower the exchange value of its currency below the PPP. Similarly, the expectation of domestic deflation—for example, in order to restore a prewar gold parity of the currency—may lead to a currency overvaluation.
5. While the PPP is the primary determinant of the equilibrium exchange rate, a secondary influence is the pattern of relative prices in each country (domestic and foreign).
6. The equilibrium exchange rate is also affected by structural variables in the countries, that is, by the demand and supply of factors of production and by production functions.
7. Changes in relative prices within a country are an indicator of real changes in the economy from a base period, and so involve a divergence between relative PPP and the exchange rate. In particular, if its export prices increase more than prices in general, a country's currency will become undervalued with respect to the PPP.
8. Long-term capital movements can drive the exchange rate away from the PPP. For example, a net long-term capital outflow may depress a country's currency below the PPP. This effect can occur only until the transfer of financial capital is fully realized in real terms, that is, in a corresponding change (in this case, an improvement) in the country's current account.
9. A private short-term capital outflow induced by the desire to evade taxation at home will cause an undervaluation of the country's currency in relation to the PPP.
10. There may be a situation in which a country cannot readily obtain capital inflows to finance a balance-of-trade deficit, and yet the commodity imports are price-inelastic (perhaps because imports of necessities are involved). In this circumstance, both the

private sector and government will bid up the price of foreign exchange above the PPP by demanding a specified amount of foreign currency irrespective of price. Here the short-term capital outflow depressing the exchange value of the domestic currency is both private and official in nature.

11. The case of a managed float is recognized. The domestic government, possibly supported by credits from abroad, can intervene in the foreign-exchange market and peg the exchange value of the country's currency above the PPP.

2.10 Policy Implications

Cassel draws a number of policy implications from his PPP analysis. These guides to government policy may be summarized as follows.

1. The PPP is the ideal rate of exchange from the standpoint of good international relations. For example, if a country's currency is undervalued with respect to the PPP, its exports are effectively subsidized and its imports hindered, much to the annoyance of traders abroad.
2. Direct measures to improve a country's trade balance are an ineffective means of increasing the exchange value of a country's currency. Given stable monetary conditions abroad, the external value of a country's currency will be largely determined by its internal value.
3. Similarly, exchange control should not be used to counter adverse speculation against a country's currency. First of all, speculation has little influence on the exchange rate. Second, exchange control can have deleterious effects and is ineffective insofar as it attempts to prevent a falling internal value of the country's currency from manifesting itself on the foreign-exchange market.
4. Writing in the early and mid-1920s, Cassel warns against countries returning to the gold standard at the prewar parities (referring not to the PPP but to the rate of exchange or mint parity). If countries are to revert to a gold standard, they should do so by fixing the exchange rate (or mint parity) at the level of the current PPP. Otherwise, for countries that have experienced large-scale increases in their price level since 1913, a severe deflationary process will be required to drive the price level down to support an exchange rate set at the prewar parity. This deflation will involve a substantial decline

in output and serious unemployment. Again, Cassel is recognizing real effects of a monetary change, in this case, a severely restrictive monetary policy.

5. If the gold standard is re-established, the spread between buying and selling points should not be reduced. A narrower band would restrict the scope of the adjustment mechanism that counteracts deviations of the exchange rate from the PPP.
6. Instability in exchange rates and in internal values of currencies should be avoided. To stabilize their exchange rates, each country must select an internal value for its currency, that is, a particular price level, and support it by suitable control of the money supply.
7. Because purchasing power parities represent equilibrium exchange rates, they should be computed and placed in the public domain regularly on a monthly basis. To this end, suitable price indexes measuring the extent of inflation in different countries and calculated on a uniform basis should be provided.

In tribute to Cassel's great accomplishment of making the PPP theory fully operational, this chapter closes with his plea for more and better data for use in applying the theory.

Notes

1. The first to do so was A. C. Pigou (1922), who used the terms "positive" and "comparative." These terms came to be replaced with "absolute" and "relative," respectively.
2. Of course, not all participants in this discussion can be readily classified into one group or the other. More interesting, only one writer apparently moved from one camp to the other. John Maynard Keynes was editor, later co-editor, of the *Economic Journal*, at the time that Cassel's first writings on PPP were published, principally in that journal. As editor, Keynes presumably played an important role in accepting Cassel's articles for publication. He also commented favorably on Cassel's theory, in two editorial notes—one appended to Cassel's first article (Cassel, 1916a), the other independently written by Keynes (1919) in the same issue as Cassel's fifth article (1919)—and in Keynes' *Tract on Monetary Reform* (1923). By the time of his *Treatise* (1930), however, Keynes had become a severe critic of PPP theory.
3. See Officer (1982, chapter 4).
4. See Officer (1982, chapter 6).

5. The f function relates the long-run equilibrium exchange rate (number of units of foreign currency per unit of domestic currency) to the PPP (foreign-country/domestic-country price-level ratio) and other variables *plus* an error term. The tendency is for the long-run equilibrium exchange rate to equal the PPP.

   The g function relates the short-run equilibrium exchange rate to the long-run equilibrium exchange rate and other variables *plus* an error term. Again, the tendency is for the short-run equilibrium exchange rate to equal the long-run equilibrium exchange rate. The logical definition of the *short-run* equilibrium exchange rate is the rate that would exist under a freely floating exchange-rate system. The *long-run* equilibrium exchange rate is discussed in Officer (1982, chapter 2, Sect. 2).

   The h function, derived as g · f, relates the short-run equilibrium exchange rate to PPP and other variables *plus* an error term. Thus the short-run equilibrium exchange rate tends to equal the PPP.

   Therefore PPP theory asserts that the exchange rate has a *tendency* to equal the PPP. This does *not* mean that PPP theory in general—and Cassel's theory in particular—has strict-equality form (long-run equilibrium exchange rate exactly equals PPP, short-run equilibrium exchange rate exactly equals long-run equilibrium exchange rate, short-run equilibrium exchange rate exactly equals PPP)—see Sects. 2.8–2.9.
6. The term is not used by Cassel; he refers to "cases where inflation proceeds with great violence and is so irregular that its progress cannot be foreseen" (Cassel 1924, p. 69). The experiences of Germany and Austria after World War I are used as empirical examples.
7. This is the "second concept" of relative PPP exposited in Officer (1982, chapter 2).
8. Typically in economic analysis, relationships—whether functional or equilibrium—are presented void of an explicit error term. (This is not true of econometric work, of course.) The question of whether a random-error term is implicitly incorporated in the relationship is never raised, because an affirmative answer is so obvious! It is strange that, of all economic theories, only PPP has been attacked for established practice.
9. Summaries of Cassel's acknowledged non-PPP influences on the exchange rate are also provided by Angell (1926), Bunting (1939), Sadie (1948), Holmes (1967), and Myhrman (1976).

References

Angell, James W. *The Theory of International Prices: History, Criticism and Restatement*. Cambridge, Mass.: Harvard University Press, 1926.

Bunting, Frederick H. "The Purchasing Power Parity Theory Reexamined." *Southern Economic Journal* 5 (January 1939): 282–301.
Cassel, Gustav. "The Present Situation of the Foreign Exchanges." *Economic Journal* 26 (March 1916a): 62–65.
———. "The Present Situation of the Foreign Exchanges." *Economic Journal* 26 (September 1916b): 319–323.
———. *Germany's Economic Power of Resistance*. New York: Jackson, 1916c.
———. "The Depreciation of Gold." *Economic Journal* 27 (September 1917): 346–354.
———. "Abnormal Deviations in International Exchanges." *Economic Journal* 28 (December 1918): 413–415.
———. "The Depreciation of the German Mark." *Economic Journal* 29 (December 1919): 492–496.
———. "Further Observations on the World's Monetary Problem." *Economic Journal* 30 (March 1920a): 39–45.
———. "Some Leading Propositions for an International Discussion of the World's Monetary Problem." *Annals of the American Academy* 89 (May 1920b): 259–267.
———. *The World's Monetary Problems*. London: Constable, 1921.
———. *Money and Foreign Exchange After 1914*. London: Constable, 1922.
———. "The Problem of the Undervalued Currencies." Skandinaviska Kreditaktiebolaget *Quarterly Report* (January 1923a): 1–5.
———. "The Restoration of the Gold Standard." *Economica* 3 (November 1923b): 171–185.
———. "The Purchasing-Power Parity." Skandinaviska Kreditaktiebolaget *Quarterly Report* (October 1924): 68–70.
———. *Fundamental Thoughts in Economics*. New York: Harcourt, Brace, 1925a.
———. "Sweden's Experiences of the Gold Standard." Skandinaviska Kreditaktiebolaget *Quarterly Report* (January 1925b): 5–9.
———. "Rates of Exchange and Purchasing-Power Parity." Skandinaviska Kreditaktiebolaget *Quarterly Report* (April 1925c): 17–21.
———. "Some Recent Experiences of Monetary Policy." Skandinaviska Kreditaktiebolaget *Quarterly Report* (October 1925d): 55–58.
———. "The Restoration of Gold as a Universal Monetary Standard." In *European Currency and Finance*. Edited by John Parke Young. Commission of Gold and Silver Inquiry, United States Senate, *Foreign Currency and Exchange Investigation*, Serial 9, Volume I. Washington: Government Printing Office, 1925e, pp. 205–206.
———. "Exchanges, Foreign." In *Encyclopedia Britannica*. Supplementary Volume I, 13th edition, 1926, pp. 1086–1089.
———. *Post-War Monetary Stabilization*. New York: Columbia University Press, 1928a.

———. "The International Movements of Capital." In *Foreign Investments.* Chicago: University of Chicago Press, 1928b, pp. 1–93.

———. "The Treatment of Price Problems." *Economic Journal* 38 (December 1928c): 589–592.

———. *The Crisis in the World's Monetary System*. Oxford: Oxford University Press, 1932a.

———. *The Theory of Social Economy.* New York: Harcourt, Brace, revised edition, 1932b.

———. *The Downfall of the Gold Standard*. Oxford: Oxford University Press, 1936.

Einzig, Paul. *The History of Foreign Exchange*, 2nd edn. London: Macmillan, 1970.

Ellis, Howard S. *German Monetary Theory, 1905–1933.* Cambridge, Mass.: Harvard University Press, 1934.

Holmes, James M. "The Purchasing-Power-Parity Theory: In Defense of Gustav Cassel as a Modern Theorist." *Journal of Political Economy* 75 (October 1967): 686–695.

Katseli-Papaefstratiou, Louka T. *The Reemergence of the Purchasing Power Parity Doctrine in the 1970s.* Special Papers in International Economics No. 13. International Finance Section, Princeton University, 1979.

Keynes, John Maynard. "Official Figures Relating to Inflation." *Economic Journal* 29 (December 1919): 504–506.

———. *A Tract on Monetary Reform*. London: Macmillan, 1923.

Meinich, Per. *A Monetary General Equilibrium Theory for an International Economy.* Oslo: The Norwegian Research Council for Science and the Humanities, 1968.

Myhrman, Johan. "Experiences of Flexible Exchange Rates in Earlier Periods: Theories, Evidence and a New View." *Scandinavian Journal of Economics* 78 (No. 2, 1976): 169–196.

Officer, Lawrence H. *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance*. Greenwich, Conn.: JAI Press, 1982.

Pigou, A. C. "The Foreign Exchanges." *Quarterly Journal of Economics* 37 (November 1922): 52–74.

Sadie, J. L. "Further Observations on Foreign Exchange Rates." *South African Journal of Economics* 16 (June 1948): 194–201.

Samuelson, Paul A. "Theoretical Notes on Trade Problems." *Review of Economics and Statistics* 46 (May 1964): 145–154.

Schumpeter, Joseph A. *History of Economic Analysis.* New York: Oxford University Press, 1954.

CHAPTER 3

Purchasing Power Parity in Economic History

Republished with permission of Wiley, from *Handbook of Exchange Rates*. Edited by Jessica James, Ian W. Marsh, and Lucio Sarno, 2012, pp. 161–187; permission conveyed through Copyright Clearance Center, Inc.

3.1 Introduction

This essay, which draws on Officer (2006), surveys the application of purchasing power parity (PPP) to historical experiences. To be considered in the historical domain and therefore included in this essay, a study's time period must fully antedate the year 1940. This arbitrary bar means that World War II and the Bretton Woods system are "post-history." The many fixed- and floating-exchange-rate episodes before Bretton Woods enable a logical ordering of the essay. The literature is surveyed according to historical periods, with each period delineated according to exchange-rate regime or regimes.

Section 3.2 categorizes PPP theories, while Sect. 3.3 presents applications of PPP to the premodern period. Section 3.4 outlines the various methods of testing the theory, and Sect. 3.5 discusses the all-important price concept in PPP. Tests of the theory for the modern period (eighteenth century to 1940) are covered in Sect. 3.6. PPP analysis of the United States return to the gold standard is discussed in Sect. 3.7. Section 3.8 looks at actual situations (in the interwar period)

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_3

in which PPP was applied to determine a new exchange rate. Concluding comments are in Sect. 3.9.

3.2 Categorization of Purchasing-Power-Parity Theories

This essay interprets PPP theory broadly.[1] Consider the variables P (domestic price index), P* (foreign price index), E (nominal exchange rate), R (PPP), and Q (real exchange rate), where R = P*/P and Q = E · R. E is defined as the number of units of domestic currency per unit of foreign currency, but may alternatively be expressed as an index number; R may be re expressed as an index number; and Q is always dimensionless.

Any PPP theory can be represented by the implicit function G(E, P, P*, X), where X is a vector of variables that can include (i) E, P, P* in earlier periods and (ii) additional variables in the current period and in earlier periods. For a specific G function to be considered a PPP theory, it is necessary that certain minimum requirements be satisfied. First, the G equation must be solvable in terms of E: E = g(P, P*, X). The E that results from solving the G function may be the actual exchange rate in the current period, the equilibrium exchange rate in the current period, or the long-run equilibrium exchange rate. Second, partial derivatives must have sign consistent with PPP theory: $\partial E/\partial P > 0$, $\partial E/\partial P^{*} < 0$.

Inclusion of (ii) variables other than E, P, and P* in G results in an "augmented PPP theory"' (the term suggested in Officer, 1982, p. 188). Is an augmented PPP theory legitimately classified within the domain of PPP? Reasonable scholars may differ on this point, but a sensible statement is as follows: The greater the importance of R (or P and P^* individually) relative to the other determinants of E, the more clearly the augmented theory is in the PPP rubric.

The variables E, P, and P* may enter in several ways. The general G function involves a trivariable theory: E, P, and P* entering as separate variables. A bivariable theory combines two of these variables; generally, R replaces P and P* in G. The theory is univariable if Q then replaces R and E. A theory has the property of "symmetry" if there are identical magnitude effects of the domestic and foreign price levels on the exchange rate ($\partial E/\partial P/\partial E/\partial P^{*} = -1$) and "proportionality" if that identical magnitude is unity. The theory in terms of the real exchange rate is proportional if $\partial Q/\partial R$ is a constant. Linear or log-linear G functions can yield these properties.

Univariable, bivariable, and trivariable theories may or may not have the property of "exclusiveness" (the term coined by Edison, 1987, p. 378). The G function involves exclusivity (perhaps a better term) if the X vector does not include any variables other than lagged E or lagged R (or lagged P and P*). Define "superexclusivity" as the absence even of the latter variables; under symmetry, the G function reduces to G(E, R) = 0. If PPP theory is univariable and super-exclusive, the G function becomes G(Q) = 0. If the G function is linear, it reduces to Q = c, where c is a constant.

The modern literature considers the R-E direction of causation irrelevant and PPP theory simply describes an equilibrium relationship among the nominal exchange rate and price levels. However, traditionally and in the historical literature, PPP theory had a causal component, implicit or explicit: prices determine the exchange rate (say, R determines E). In terms of the real exchange rate: if Q is shocked out of equilibrium, R (rather than E) changes to restore equilibrium.

What modern economists cannot legitimately deny is that PPP is a monetarist theory and, as such, asserts that, at least in the long run, the nominal exchange rate (E—a monetary variable) can be affected only by monetary variables, such as PPP (R). In contrast, the real exchange rate (Q) is a real variable; correspondingly, in the long run only real variables can affect it.

3.3 Historical Application of PPP: Premodern Periods

The "premodern" period denotes human history before the eighteenth century. For the premodern period, the only use of PPP is to assess the extent of the integration of the domestic economy with foreign economies. The closer PPP is to fulfillment, the greater the integration. The best way of using PPP to determine the amount of integration of economies is to test PPP theory statistically; but this method is not possible for premodern economies, because of the lack of data. Alternatively, one could observe either individual-commodity price differences, in domestic currency, at home and abroad, or exchange rates and domestic prices. This technique is usable, even in the absence of recorded price series; for (i) contemporary authors may have written of the price differences, or (ii) inferences on price differences may be made by modern scholars on the basis of other information.

Premodern economies are characterized by monetary systems in the realm of a coin standard. Except in China, paper standards were unknown until the eighteenth century. In fact, again except in China, paper money did not even exist until toward the end of the premodern period. Exchange rates were "fixed" at mint parities. According to Einzig (1970, p. 71), foreign-exchange transactions were almost entirely coin-for-coin until the thirteenth century, when bills of exchange became dominant. The "fixity" of exchange rates was not absolute, in two respects. First, mint parities were responsive to depreciation and debasement of coins. Second, specie-point spreads were much wider than in modern times.

3.3.1 Ancient Period

The few scholars who have examined the issue are unanimous that there was not even a tendency for PPP theory to be fulfilled in ancient times.[2] In particular, there are four reasons why PPP theory did not apply to the Roman Empire. First, Roman imports were luxuries, such as valuable furs, amber, carpets, silk, precious stones, and aromatics, which were purchased only by the rich. They were income, rather than price, determined. The own-price elasticity of demand for imports was effectively zero. Second, imports were not produced in the Empire, and had no close domestic substitutes. The cross-price elasticity of demand for imports with respect to domestic commodities was also zero, or close to it.

Third, there were tremendous price differences between Rome and its trading partners.[3] Such price differences are suggestive of arbitrage imperfections, which took two forms: high transportation and commission charges, inherent in the state of transportation and communication technology; high profit margins and risk premiums. Fourth, trade in nonmonetary commodities between the Roman Empire and Asia was largely one-sided. Rome imported luxury goods; but exported nonmonetary commodities to the Far East only in small amounts, because Far Eastern countries had little demand for wine, oils, wool manufactures, and leather manufactures—which were the Empire's principal exportables. Rome did have an abundant exportable that was in demand in the East: silver. To the extent that silver exports were in the form of coin, Rome thereby financed its "balance-of-payments deficit on commodity account." To the extent that the silver was bullion, it can be considered a normal commodity export, reducing the one-sided character of Roman-Eastern trade.

In sum, the economy of the Roman Empire may have been well integrated within itself, but certainly not with the outside world.

3.3.2 Medieval Period

Einzig (1970, p. 99) states that "in the Middle Ages…exports and imports were largely inelastic and unresponsive to changes in prices or exchanges." Officer (1982, p. 28) suggests reasons for this inelasticity. First, the feudal economy, with its self-sufficiency nature and structured society, had purely luxury imports, as in Roman times. Second, general contraction of trade occurred, both within Europe (the former Roman Empire) and between Europe and Asia. The result was again inapplicability of PPP and the lack of integration between economies, even within Europe.

Over time, trade expanded and parts of Europe coalesced into sovereign regions. Trade among these entities took place, and foreign-exchange markets developed. What used to be intra-Empire trade became foreign trade, with distances and risks less than they had been for Empire-Asian trade. It is reasonable to presume that commodity arbitrage gradually became less imperfect over time, and that there was an increasing tendency toward PPP, as economies became more integrated.

3.3.3 Sixteenth-Century Spain

Spanish scholars of the Salamanca School originated the PPP theory. In their environment, PPP was an indicator not only of integration of the Spanish and outside economies but also of the importance of monetary influences on the exchange rate. The Salamancans made the following empirical observations: (i) Spain had received large inflows of gold and silver from the New World; (ii) the Spanish money stock increased; (iii) the Spanish price level also increased; and (iv) exchange rates had become unfavorable to Spain. Spain (along with England and the rest of Europe) was on a metallic standard. Therefore, what an unfavorable movement in exchange rates meant was a movement in current exchange rates away from mint parities in the direction of specie-export points. This was a lower exchange value for Spanish coin. Thus PPP as an equilibrium theory was fulfilled, at least in an approximate sense. The causal PPP theory was also satisfied, with the obvious causal chain (i) → (ii) → (iii) → (iv).

3.4 Techniques of Testing PPP Theory in Economic-History Literature

3.4.1 Comparative-Static Computation

Let E_0 denote E in period 0. An obvious test of PPP theory is to measure P and P* as index numbers with value unity in base period 0 and compute $V = (P/P^*) \cdot E_0$ for either one period, a few discontinuous periods, or a continuous sequence of periods. The computed V are then compared with the corresponding values of E, in a table or graph. Alternatively, $(E/E_0)/(P/P^*)$ is compared with unity. In either case, the closer the computed value to the norm, PPP-predicted, value, the closer is PPP theory to fulfillment. Any noticeable divergences are then explained in terms of non-PPP influences on the exchange rate (augmented PPP theory). One can allow for a lagged effect of R on E. Further, investigations of lead-lag relationships are used to test the PPP-postulated direction of causality, from prices to the exchange rate. This entire approach has the "advantage" of lying outside formal statistical analysis.

3.4.2 Regression Analysis

The use of regression analysis was a natural development in testing PPP theory. For example (using lower-case letters to denote logarithms), e is regressed on p and p* or on r^{-1}; q is regressed on a constant. Properties such as symmetry and. proportionality can be readily tested in terms of elasticities.

3.4.3 Testing for Causality

The PPP relationship tested can either be an equilibrium relationship or a causal relationship, each being tested directly. A hybrid test involves an equilibrium relationship tested via an imposed causal relationship. The causal direction isindicated by the direction of minimization of the sum of squared errors in regression. Generally, minimization is in the direction of the exchange rate. Some scholars believe that this direction of minimization is applicable only to a floating exchange rate. Under a fixed exchange rate, the "dependent variable" is p and the "independent variable" $e \cdot p^*$: the domestic price index is determined by the foreign price index expressed in domestic currency, the "world" price index governing the domestic price. However, Gustav Cassel and other traditionalists considered PPP as the principal determinant of the exchange rate under both a floating exchange rate and the gold standard.[4]

In modern work, testing for causality involves a peculiar definition of causality: Granger causality, a forecasting concept. If r^{-1} aids in the forecasting of e beyond lagged values of e, then r^{-1} is said to Granger-cause e. Only preceding values of r^{-1} can assist in the forecast; and they can only assist: lagged values of e are also in the forecast equation.

3.4.4 Nonstationarity and Spurious Regression

It came to be realized that regression analysis ignored crucial time-series properties of variables, so that the results and conclusions could be meaningless or spurious. The comparative-static and regression studies of PPP implicitly assume stationarity of the PPP, nominal-exchange-rate, and real-exchange-rate series that are utilized. A stationary series has a constant and finite mean, a constant and finite variance, and covariances that are constant for a given time interval between the observations.

A stationary series has several desirable properties. (i) Computation of the parameters (mean, variance, autocorrelations) of the series is readily accomplished from sample data. (ii) The series exhibits "mean reversion": at least in the long run, the series returns to its mean, the equilibrium value of the series. Deviations of the series from its mean are only temporary. Shocks to the series have only temporary effect. (iii) There is the possibility (although not the necessity) of "short memory," implying a relatively fast reversion to the mean, after any disturbance. (iv) There is no statistical reason why regressions or correlations involving only stationary series would be spurious. The legitimacy of standardized tests for significance stands.

If d is an integer (the usual assumption), the "order of integration" (d) of a series is the minimum number of times that the series must be differenced to achieve stationarity. Traditionally, economists have an either-or viewpoint of stationarity: a series is stationary as it stands ($d = 0$) or needs to be differenced once to achieve stationarity ($d = 1$). And the conventional wisdom has been that most economic series are in the latter category. Nonstationary series ($d \geqq 1$) have disadvantages. (i) They lack a constant mean; or, if they have such a mean, have a nonconstant or even infinite variance. Estimation of parameters of the series cannot be readily effected. (ii) Mean reversion is not present. There is no mean to which to revert; or deviations from a mean persist. (iii) Memory is infinitely long; the "stochastic trend" emanating from disturbances adds new terms without reducing the impact of existing terms. (iv) Regressions and correlations can be spurious with even one of the variables nonstationary, and standard tests of significance give misleading results.

It is now realized that a series can be "fractionally integrated" ($0 < d < 1$). Such series are mean reverting, but have long memory. It takes a long time for the effects of disturbances to die out. The critical value for d is 1/2. For $d < 1/2$, the series is stationary; for $1/2 \leqq d < 1$, the series is nonstationary, because the variance of the series is infinite. In the latter case, stationarity is obtained by first-differencing the series. For correlations or regressions involving two fractionally integrated variables, say of orders d_1 and d_2, spurious results occur for $d_1 + d_2 \geqq 1/2$. Of course, if the variables are fractionally cointegrated, then the results would not be spurious.

3.4.5 Testing for Stationarity

Modern univariate testing of PPP involves testing for the stationarity (order of integration) of the real exchange rate (q) as distinct from testing for a constant mean of q (as was formerly done). Only if the series is stationary is there an equilibrium value of the real exchange rate to which the actual value reverts in the long run. Deviations of the actual from the equilibrium (mean) real rate do occur, but they eventually disappear. "Eventually" is not necessarily good enough. For the stationary series, a PPP-shock half-life (the length of time needed for the original deviation of actual q from equilibrium q to be halved) is a crucial statistic. The longer this half-life, the less is PPP theory supported. These tests of PPP are pleasing, because they treat the theory as applicable only to the long run; but the tests are also displeasing, because (i) a PPP-determined mean value is not imposed and (ii) symmetry and proportionality are ignored. In other words, an extremely weak interpretation of PPP is tested.

While a series can be made stationary by first-differencing, it is also possible that a transformation short of first-differencing might work, for example, taking logarithms of the variables. Also, including the lagged dependent variable as an explanatory variable might be sufficient. While early studies of PPP paid no attention to stationarity (and, as seen in

Sect. 3.6, these are the bulk of the studies in the economic-history vein), some did make an adjustment—perhaps inadvertently or for some other reason (usually hypothesis specification).

3.4.6 Cointegration Analysis

Even if variables are tested and nonstationarity is found, one should not proceed with correcting for spurious correlation; for the variables might be cointegrated, that is, a linear combination of the variables is stationary. Cointegration analysis is admirably suited for PPP bivariate or trivariate testing. The cointegration model involves the PPP relationship with zero error, as holding in the long run: it is the long-run equilibrium. Deviations from that relationship occur in the short run, and an error-correction process returns the variables to the long-run relationship: there is mean reversion. Differing speed of adjustment for the exchange rate and for price (or price ratio) is embedded in the error-correction process. Symmetry and proportionality are not imposed, and can be tested. General-to-specific modeling, in which restrictive models are nested within more-general models, can be used here, as with conventional regression analysis. The causal aspect of PPP can also be tested, via the speed-of-adjustment coefficients. Fractional cointegration, involving fractionally integrated variables, is also possible, although rarely performed in the PPP literature.

3.5 Price Variable in PPP Computations

Crucial to empirical use and testing of PPP is the price concept, and many price concepts have been used in PPP computations. Ranging from most justifiable to least justifiable, they are as follows (with symbols): GDP deflator (PGDP), GNP deflator (PGNP), consumption deflator (PCONS), retail price index [incorporating consumer price index and cost-of-living index] (RPI), wholesale price index (WPI), export price index (XPI), wage-rate index (WI), component indexes or subindexes of WPI or RPI, and prices of individual commodities.

The bar separating PPP-legitimate price measures is drawn between WPI and XPI. So only results based on PGDP, PGNP, PCONS, RPI, and WPI are included in the survey. PGDP and PGNP have three justifications. First, as stated by Cassel (1928, p. 33), PPP relates to the internal value of currencies, and therefore should be “measured only by general index figures representing as far as possible the whole mass of commodities marketed in the country." PGDP and PGNP fit this criterion better than any other price index. Second, PPP is a macroeconomic theory, and therefore necessitates the usual macro-price concept, PGDP or PGNP, with the former marginally preferred, because it measures the price of production within the country. Third, to the extent that PPP is justified by arbitrage and substitutability of commodities in production and consumption (broadly construed), the price concept underlying PPP should be as broad as possible, again leading to PGDP or PGNP.

Other things being equal, one would like to place WPI below the separation bar while leaving RPI above the bar for two reasons. First, a PPP computed from traded-goods prices alone is close to a truism.[5] Because (i) in any given country, the WPI is heavily weighted with tradables and, in particular, excludes all services and (ii) across countries, arbitrage directly equates prices of tradables (up to transactions costs, including tariffs and transportation charges), a PPP computed from WPIs comes close to making PPP theory a truism. Therefore, WPI biases result in favor of the hypothesis that PPP theory holds, and therefore that the domestic economy is well integrated with the foreign economy. In contrast, the RPI consists of nontradables (services) as well as tradables. Also, the weighting pattern of the WPI need not bear a close relationship with the production-weighted (i.e., GDP weighting pattern) of the economy. The WPI incorporates considerable, but unknown, double-counting and even multiple-counting. In contrast, the RPI has a logical weighting pattern. Yet, unfortunately the most widely used price measure in PPP studies is the WPI. For many historical periods, the WPI is the only, or at least the most comprehensive, price index available. Even when alternative indexes exist, researchers often select the WPI. So, on grounds of expediency the WPI just makes the bar.

The XPI is totally composed of tradables. After the joyless decision to include the WPI above the bar, it gives one a certain pleasure to place the XPI below the bar. Also, price measures of individual commodities are excluded, because of their lack of comprehensiveness. WIs are excluded, primarily because PPP theory (and its justifications) pertains to prices of commodities rather than of factors of production. A second reason to exclude WI is the opposite justification for excluding XPI (and only reluctantly including WPI). In contrast to measures heavily weighted with tradables, that can move almost automatically in correspondence with the exchange rate, WI is typically the price of an entity heavily nontradable, and can move *too slowly* relative to the exchange rate.

3.6 Modern Period: Testing of PPP

The principal use of PPP in historical research of the modern period (eighteenth century onward) is testing the validity of the theory. Although almost all investigators test the theory for its own sake, in effect, the degree of integration of the domestic with the foreign economy is assessed. Most, but not all, studies pertain to periods of a floating as distinct from fixed exchange rate.

3.6.1 Early North America

The earliest date of any PPP testing in this survey is the U.S. colonial period, and all authors paid attention to the stationarity issue. Bordo and Marcotte (1987) found that PPP holds under the South Carolina adjustably fixed exchange rate and proportionality could not be rejected. Choudhry and Luintel (2001) examined Pennsylvania under a floating exchange rate, and PPP results are mixed.

Bernholz (2003) examined the period of the 13 colonies in rebellion, during.which Congress issued Continental currency. During this paper standard and floating exchange rate, the price of specie (silver coin—representing the exchange rate) did not increase as much as the price index. Bernholz's explanation is the war-inflicted damage on production (supply) of goods and the British blockade,which reduced the value of specie (the currency used in payment for imports).

Grubb (2003, 2005, 2010) compared properties of the real exchange rate for six American colonies (later U.S. states) and Lower Canada in 1748–1775 (colonial period) versus 1796–1811 (Constitution period), with Lower Canada serving as a control. In the colonial period, only Massachusetts and Lower Canada were clearly on specie standards and fixed exchange rates. In the Constitution period, all U.S. states were on a fixed exchange rate—that of the U.S. dollar—by default. Lower Canada was now on a floating exchange rate, by virtue of Britain abandoning the gold standard in 1797. In general, PPP held; but half-lives to shocks were lower in the colonial than Constitution period. As Grubb (2010, p. 141) stated: "Market integration as measured by PPP during 1796–1811 was not superior to market integration during the 1748–1755 colonial period."

3.6.2 Bullionist Periods

A "bullionist period" in economic history has both an empirical and an intellectual characteristic. Empirically, a bullionist period involved a paper standard and floating exchange rate that temporarily interrupted a specie standard and fixed exchange rate. Intellectually, a bullionist period carried with it a "bullionist controversy" regarding the ruling macroeconomic model of the economy. In modern terminology, the competing models are monetarist and nonmonetarist. In particular, "bullionists" were monetarists, and generally exposited a PPP theory of the exchange rate.

Two bullionist experiences that have been subject to PPP testing are the Swedish bullionist period (1745–1776) and the Engish bullionist period (1797–1821). The latter is customarily called the Bank Restriction Period, because the Bank of England's obligation to pay cash (gold) for its note issues was restricted. It may be noted that, while paper money originated in China, banknotes were first issued in Sweden. The Swedish bullionist period began with the paper daler made inconvertible into copper bullion.

Three authors investigated PPP for both episodes. Eagly (1968, 1971) noted increases in the price level and exchange rate in terms of banknotes. Myhrman used growth rates and found positive evidence for PPP. Bernholz, Gärtner, and Heri [hereafter Bernholz et al.] (1985) applied a univariate technique to various floating-rate episodes and found that PPP was violated in the short run but held in the long run; however, they paid no attention to stationarity. Bernholz's (1982, 2003) results for Sweden are consistent with those of Bernholz et al.

Turning to authors who examined the Bank Restriction Period exclusively, Angell (1926, p. 484) found no relationship between the British price index and exchange rate. Nachane and Hatekar (1995) rejected cointegration of the British price index and exchange rate. Also, they could not reject that price does not Granger-cause the exchange rate. Their use of the exchange rate on Paris is contrary to other researchers;

economic historians generally view the Hamburg exchange as more representative than Paris during the French Revolutionary and Napoleonic Wars. The Nachane and Hatekar observation period extends to 1838, which, with annual data, increases the sample size. However, the effect is a mixture of a "paper standard, floating exchange rate" with a "gold standard, fixed exchange rate."

In contrast, Officer (2000) used quarterly data and limited the sample to the Bank Restriction Period. Another difference is that Officer engaged in multivariate testing, with Bank of England notes, the price of wheat, and external military expenditure as variables in addition to the general price index and exchange rate. While Nachane and Hatekar also employed multiple variables, their testing was entirely bivariate. However, Officer's results regarding PPP were negative, and essentially the same as those of Nachane and Hatekar.

3.6.3 Floating Rates—Second-Half of Nineteenth Century

The greenback-period episode of a paper standard and floating exchange rate encompasses the full years 1862–1878, and has received considerable attention in the literature. Graham (1922) did not mention PPP, and in fact made no formal judgment on the validity of the theory. He stated that the principal determinants of the exchange rate were (i) expectations regarding a Northern victory, during the Civil War, and (ii) net capital inflow, in the postbellum period. However, it is clear from his computations and the context that these influences were secondary, and came into play given the effect of commodity prices on the exchange rate. This was the position also of later authors who offered an augmented PPP theory for the greenback period: Kindahl (1961) and Friedman and Schwartz (1963). None of these authors paid attention to nonstationarity; nor did Farag and Ott (1964) and Thompson (1972). Therefore, the generally positive results of all these early writers were questionable; and the regression analyses of Farag and Ott (1964) and Thompson (1972) could be particularly misleading, as the estimation technique is ordinary least-squares.

Officer (1981) provided some innovations to PPP investigation of the greenback period. Instead of representing the dollar-sterling exchange rate by the price of gold, he constructed a "true" exchange-rate series (the inverse of the dollar-sterling rate) as the ratio of the gold-dollar price of the greenback to the gold-dollar price of the pound; and his price concept is the GNP deflator rather than the WPI. Also, Officer, in effect, corrected for nonstationarity in his regressions. So his positive results for PPP lend credibility to the results of the earlier studies.

Bernholz et al. (1985) obtained their usual qualitative result of short-run PPP violated, long-run return to PPP. Enders (1989) is the only author other than Officer to have addressed nonstationarity, and he trumped Officer because his attention is deliberate rather than inadvertent. However, Enders' results were mixed. Nonstationarity ($d = 1$) could not be rejected for the real exchange rate, while the U.S. and British price indexes (the latter expressed in dollars) were found to be cointegrated. The former result is unfavorable to PPP, the latter is supportive.

Austria was on a paper standard and floating exchange rate from the mid-nineteenth century, when it left the silver standard, until 1892, when. Austria-Hungary joined the gold standard. Especially interesting is the subperiod 1879–1892, during which the paper gulden was worth more than its legal metallic content. The PPP aspect of the Austrian experience was investigated by Yeager (1969) and Myhrman (1976), who reprinted Yeager's graph of the exchange rate and PPP. Yeager (1969) computed correlation coefficients of the exchange rate and PPP, for the variables in percentage-change form, which could produce stationarity, and the results can reasonably be construed as positive evidence for PPP. Myhrman (1976, p. 190) commented that "both prices and the exchange rate were rather stable but with a rising trend." This remark is suggestive of a possible trend-stationarity characteristic of the variables.

3.6.4 Classic Metallic Standards

With the United States back on the gold standard in 1879 and Britain on gold since the end of the Bank Restriction Period in 1821, it is natural to examine PPP for these two countries in the context of the fixed exchange rate of the gold standard. Enders (1989) offered an identical analysis as for the greenback period. For this period, not only was there cointegration in a bivariate model, but also now nonstationarity of the real exchange rate was rejected—supportive of PPP. In contrast, Grilli and Kaminsky (1991) could not reject nonstationarity in the real exchange rate, destructive of PPP.

Catão and Solomou (2005) investigated real-effective exchange rates for three groups of countries: the gold-standard core group, countries on a silver standard for at least part of their time period (1871–1913), and countries on an inconvertible paper standard for a substantial part of the period. They did not formally test for nonstationarity of the real exchange rate of the countries; rather they accepted stationarity, based on estimated autoregressive coefficients uniformly below unity. Nevertheless, their results are unfavorable to PPP; for they found "large and protracted real exchange rate fluctuations" (Catão and Solomou, 2005, p. 1265).

Considering both the gold standard (1880–1914) and interwar period (1921–1940) for the United Kingdom and Canada versus the United States, McCloskey and Zecher (1984) found that PPP regression forecast errors were not related to U.S. money-market disequilibria—supportive of PPP. Their use of the GNP deflator is to be commended. Hasan (2004) examined PPP for silver-standard India. Among other techniques, he applied fractional-integration modeling. The hypothesis of nonstationarity of the real exchange rate was rejected.

Hegwood and Papell (2002) studied Belgium, France, Germany, and the United States over 1793–1913, which encompassed episodes of the gold standard, silver standard, and paper standard. Impressive is their concept of "quasi-PPP": reversion to mean q that exhibits structural shifts rather than to a constant mean q; and half-lives of PPP deviations were short. Structural breaks were associated with economic and political events: the U.S. Civil War, dissolution of the German Confederation, coup d'état in France, and the 1840s decade of political unrest in Europe.

3.6.5 World War I

Investigations of PPP for World War I were undertaken by Cassel (1916, 1918, 1919), Heckscher (1930), Keynes (1919), and Bresciani-Turroni (1937). Gustav Cassel, the greatest expositor and propagator of PPP, naturally tested the theory first for his own country. All these studies were comparative-static in nature.

Findings were mixed, and, because there was no attention to nonstationarity, must be viewed with caution.

3.6.6 Floating Rates—1920s

A tremendous number of PPP studies pertains to the 1920s, especially the first part of that decade. There are several reasons for this concentration. First, all countries on the classical gold standard left gold during World War, resulting in floating exchange rates. During the war, there was exchange-market intervention on the part of some countries, in particular, France and the United Kingdom. However, shortly after the end of the war, almost all exchange rates became freely floating, and remained so for a substantial part of the decade. PPP theory is typically of greater interest to researchers when exchange rates are floating rather than fixed. Second, the United States, which had effectively adopted a paper standard extralegally in April 1917, returned to gold in March 1922. The dollar thus provided an anchor to other countries for a return to a fixed-rate system and for assessing the level and volatility of the exchange rate while their currencies were floating. In particular, the United States was a natural base country for PPP computations. Third, even though World War I marked the end of the nominal international economic supremacy of the United Kingdom and even though that country did not readopt the gold standard until April 1925, its traditional importance as the center country of the classical gold standard made it a natural alternative base country for floating exchange rates of other countries. Fourth, the very fact that the once central country of a metallic standard and fixed-rate system (the gold standard) was now floating made the United Kingdom a most interesting subject of PPP analysis, with the United States (the upstart other center country) as base country. Fifth, for researchers in the final quarter of the twentieth century and beyond, it was natural to compare the floating rates of the post-Bretton Woods period with the floating rates of the 1920s, in particular, from the standpoint of PPP analysis.

The U.K. floating rate of 1919–1925 has been studied by many authors.[6] While there is no consensus, the preponderance of the evidence suggests that the pound sterling in the 1920s floated in the dollar-pound exchange market in a manner consistent with the PPP theory. France had a floating exchange rate in 1919–1926, and this experience has been investigated by various authors.[7] PPP aspects of the German floating rate of 1914–1923 were considered by Bresciani-Turroni (1937), Frenkel (1976), Haberler (1936), Rogers (1929), and Bernholz et al. (1985). For each episode, results have quite different implications regarding the validity of PPP, and no general assessment of the validity of PPP can be made. Other individual-country PPP studies of the 1920s concerned the floating and fixed exchange rate of Sweden, and the floating rates of Switzerland and Greece.[8] Generally, results were negative for Sweden, mixed for Switzerland, and positive for Greece.

An impressive, two-volume, assemblage of studies of the post-World-War I monetary and exchange-rate experience of European countries is that of Young (1925a). Papers were contributed by a large number of government officials and academic economists, and many of the writers in effect test the PPP approach to the exchange rate.[9] Almost all the authors adopted a common comparative-static methodology, plotting and tabulating the exchange rate versus the PPP, with the United States as the base country. The following general findings can be discerned: (i) the exchange rate and PPP moved more or less in correspondence; (ii) there was persistent directional deviation of one of these variables from the other; (iii) the usual pattern was that the domestic currency depreciated more than indicated by PPP, that is, the currency was undervalued in the foreign-exchange market; (iv) there was closer correspondence of the exchange rate and PPP in recent years than previously, especially if exchange-rate stabilization had occurred.

Bernholz et al. (1985) applied their model to three countries in the 1920s. The usual results (PPP violated in short run, validated in long run) pertained to Hungary and Poland. Austria was an outlying case, with PPP not holding in the long run.

Many studies tested PPP in the 1920s for multiple domestic countries with the United States as the base country. They fit into two groups: those that ignored stationarity and those that addressed the issue (even if via another route).[10] Distinctive in the first group is the conclusion of Keynes (1923, pp. 101, 106), who later became a critic of PPP, that "even under such abnormal conditions as have existed since the Armistice… the Purchasing Power Parity Theory, even in its crude form, has worked passably well"; and Flux's (1924) early use of logarithms in PPP computation. In the studies that addressed stationarity, overall results are more positive than negative for PPP. There are also 1920s studies with the United Kingdom as the base country.[11] No author addressed stationarity, and yet the only positive results are those of Thomas (1972).

Finally, there are 1920s studies which do not have a base country as such. A set of "equal status" countries was selected, and the PPP between pairs of these countries was investigated. One country group consists of the United States, United Kingdom, and France; another group adds Germany to these countries.[12] All studies were conducted using modern time-series analysis, with explicit attention paid to stationarity. For the first group, results were largely mixed; for the second group, they were mostly positive. Very impressive is Michael et al. (1997), who specified a nonlinear adjustment process. They concluded that, for country pairs excluding Germany: “Despite the high degree of persistence in PPP deviations, our framework provides strong evidence of mean-reverting behavior for the real exchange rate” (Michael et al., 1997, pp. 876, 877).

3.6.7 1930s

In September 1931 the United Kingdom abandoned the gold standard for a managed float, while the United States did not leave gold until March 1933. Broadberry (1987), Whitaker and Hudgins (1977), and Grilli and Kaminsky (1991) performed PPP testing for the United Kingdom during the 1930s, with the United States as the base country. Overall, the results were negative. Graham (1935), White (1935), and Broadberry and Taylor (1992) dealt with the 1930s PPP experience of multiple countries. Overall, again results were not generally positive. The Broadberry-Taylor study is instructive. They examined all pairs in the country-group United States, United Kingdom, France, Germany, and could not reject that the real exchange rate is nonstationary ($d = 1$), a failure of PPP. Cointegration results, which pertain to equilibrium PPP, were mixed. Granger-causality tests, which address causal PPP, were also mixed. For full samples, prices never Granger-caused exchange rates, but the reverse was sometimes found. Only for subperiods of freely floating rates was there some evidence of prices Granger-causing the exchange rate.

McCloskey and Zecher (1984) found a close weekly relationship between the dollar-pound exchange rate and the U.S. WPI or RPI in 1933, with foreign prices relatively constant. The finding is associated with the depreciation of the dollar, the United States leaving the gold standard. McCloskey and Zecher (1984, p. 143) concluded: “Purchasing power parity is not a failure. On the contrary, by the standards we have examined, it is a great success.”

3.6.8 Interwar Period

Some studies treated the interwar period as a broad expanse, incorporating fixed and floating exchange rates in the same sample. Young (1938) found that there were subperiods defined by PPP and the exchange rate alternately moving together (during one subperiod) and diverging (during the next subperiod). This is not good evidence for PPP. Bunting (1939) graphed the exchange rate against PPP, with the latter alternatively lagged zero, one, two, and three periods. This is a logical way of assessing causal PPP, even though it is defective for the lack of attention to nonstationarity. Even with the lags, there were substantial deviations between the PPP and the exchange rate, and in opposite directions for France and the United Kingdom. Bunting (1939, p. 299) judged: "This is damaging statistical evidence against the purchasing power parity theory." Katano (1956, 1957) computed a number of correlation coefficients; but these are largely devoid of meaning, because of the small number of observations and the danger of spurious correlation. His most interesting result is that deviation from PPP was related to divergence from pure inflation in the countries. Bernholz et al. (1985) exhibited the usual result of PPP validated in the long run while violated in the short run.

3.6.9 Spain—Long Term

Spain was on a floating exchange rate for most of the 1870–1935 period, moving from a free to a managed float in 1931. For 1914–1920, Delaplane (1934, p. 41) used comparative-static computation to note "the wide divergence of purchasing power parity from the [exchange] rate." For the entire 1914–1933 period, his assessment of PPP was, at best, mixed: "In the light of Spanish monetary experience since 1913, one could not attribute more than a rough correspondence between purchasing power parity and exchange" (Delaplane, 1934, p. 211). Using Delaplane's data for the subperiod 1920–1929, Yeager (1976, p. 220) took a more-sanguine view of PPP: "The actual rate kept within the range of 12.5% below to 12.5% above purchasing-power parity in 82.5% of the months."

Sort et al. (2005), following Sabaté et al. (2003), examined the Spanish experience over the full 1870–1935 period. They considered q for the peseta against the British pound, French franc, and U.S. dollar. In the Hedgwood-Papell tradition, allowing for structural breaks in q enables rejection of nonstationarity. These breaks were explained via rumors of restoration of gold convertibility of the peseta in 1927, the pound abandoning the gold standard in 1931, stabilization of the peseta in 1931, and the financial instability of France after World War I. The authors concluded: "one can accept the PPP hypothesis as *a good approximation* of the behaviour of the peseta exchange rate against its main traders and investors between 1870 and 1935."

3.6.10 Guatelmala—Long Term

Schweigert (2002) investigated PPP for Guatemala for 1897–1922, during which time the country was on a floating exchange rate. The United States was the base country. The money stock was used to proxy the Guatemalan price index, for which a direct series does not exist. This representation had been adopted by Cassel and Keynes for their World War I studies, absent price data. Results were excellent for PPP. The exchange rate, U.S. price, and Guatemalan money stock were found to be cointegrated. With the coefficient of the exchange rate normalized to unity, one could not reject the hypotheses of symmetry and proportionality (coefficients of U.S. price and Guatemalan money, one and minus one, respectively).

3.7 Analysis of U.S. Return to Gold Standard in 1879

The successful PPP testing for the greenback period, on the part of Kindahl (1961) and Officer (1981), was based (wholly, for Kindahl; in part, for Officer) on real-exchange-rate computations. These authors put their computations to work to determine (i) the range of real appreciation of the greenback for successful return to the gold standard and (ii) the first year in which a successful return could occur. For (i), the technique was simply to observe the range of the real exchange rate in the postbellum period but excluding 1877–1879, which were years of unusual capital outflow. The resulting range for Kindahl was 9–27% or 8–18%; for Officer, −3 to 18%—all assuming no capital movements. If resumption was to occur at the prewar parity (as in fact did happen), then the U.S. price index could exceed the U.K. price index by a value within the specified range (with both indexes relative to base-year 1860). With capital inflow, the real exchange rate (or PPP, with no change in the nominal exchange rate) could exceed the upper limit. With capital outflow, it might have to fall below the lower limit.

To answer (ii), one approach is to find the earliest year in which the real exchange rate falls within the estimated range; but the range might be considered too broad for a confident return and maintenance of the gold standard. Consider, rather, a stronger criterion: the earliest year at which the real exchange rate reached (or almost reaches) 100—the same value as in 1860. For Kindahl, that year was 1879, when his real exchange rate was 101 and the return to gold in fact occurred. For Officer, the year was 1875, when his real exchange rate was 100 and the Resumption Act was passed. That Act specified a return to the gold standard on January 1, 1879—which in fact happened. To some historians of the period, Officer's answer would be too optimistic. For example, Friedman and Schwartz (1963, p. 48) wrote that "the act was little more than the expression of a pious hope." However, they went on to state: "Resumption might well have been successful a year or more earlier than the date set and certainly could have occurred later"—Friedman and Schwartz (1963, p. 85).

3.8 Establishment and Assessment of a Fixed Exchange Rate in Interwar Period

This section discusses the use of PPP by government in connection with the setting of a new exchange rate. In the interwar period, there were two interesting cases of PPP computations by the government in order to establish a new, or return to a former, exchange rate: United Kingdom (1925) and France (1926).

3.8.1 *United Kingdom*

The U.K. return to gold on April 28, 1925 was the (sole) case of a government predetermining the exchange rate—in this case the prewar gold parity—and using PPP to measure the amount of price-level adjustment at home or abroad required to maintain the rate. France, Czechoslovakia, and Belgium applied PPP to compute the new exchange rate, although less so in France than in the other two countries.

No doubt the U.K. experience is the most famous of all governmental applications of PPP. The floating pound had appreciated from 10% to less than 2% below parity—caused by anticipation of a return to parity, whereupon the prewar exchange value of the pound ($4.86656 per pound) was restored. There was never a question that return to the gold standard would take place, and at the prewar rate. As Sayers (1960, p. 314) commented: "The restoration of the gold standard, at a tacitly assumed rate of 4.86, was government policy throughout". Moggridge (1969, p. 14) agreed: "The Authorities had as their primary aim a return to gold…a return to the pre-war parity." There was never any choice as to the fact of return and the rate. According to Sayers (1960, p. 317), one of the advisers of Winston Churchill, Chancellor of the Exchequer, told him:

"There's no escape; you have to go back [to gold at the prewar parity]; but it will be hell."

For the authorities, the only question was timing: *when* the gold standard would be reestablished. Churchill's advisers used the WPI in their PPP computation, which Keynes criticized for virtually validating the existing exchange rate. "This led them to think that the gap to be bridged was perhaps 2 or 3 per cent"—Keynes (1931 [originally published in 1925], p. 250). So the return to gold occurred on April 28, 1925.

Both contemporary and later economists used PPP to determine the overvaluation of the pound upon re-adoption of the gold standard.[13] The earliest such computation was apparently made by Keynes himself. He contrasted the government WPI-based estimated overvaluation of 2–3%, with his own RPI-based figure of 10–12%. The former estimate was considered biased downward, the latter (in conjunction with PPP based on wages and prices of manufactures) "a much better rough-and-ready guide for this purpose…than are the index numbers of wholesale prices" (Keynes, 1931 [originally published in 1925], p. 250). However, as first pointed out by Gregory (1926), Keynes used RPI figures from the state of Massachusetts rather than the national U.S. data of the Bureau of Labor. The presumed reason, according to Gregory, was that only the former series at the time was published on a regular basis. Using the national figures, Gregory obtained results in accord with those of Churchill's advisors. Cassel (1925b, 1926) offered a WPI-based estimate slightly above that ascribed by Keynes to Churchill's advisors.

The computations of later writers used a broader array of indexes and base countries. Moggridge (1972) was the first author to employ the GNP deflator—a superior price index than the WPI and RPI—and found overvaluation to be 11%, consistent with Keynes. Moggridge (1972, p. 105) wrote: "An exchange rate at least 10 per cent lower than $4.86 would probably have been somewhat more appropriate for sterling." Dimsdale's (1981) estimates were between 1 and 14%, depending on the price index. In addition, he computed a real effective-exchange-rate for sterling versus 11 currencies, but only from 1920 and on a 1929 rather than prewar base. The work of Redmond (1984) is impressive for a wide array of alternative base countries as well as for effective-exchange-rate computations; however, as might be expected, estimates are all over the place. Matthews (1986) offered estimates based on the work of Redmond and Moggridge. Taylor's (1992) estimate of 5% overvaluation was based on an error-correction model, and is mentioned because the model is within the PPP rubric.

Certainly, the estimates of overvaluation of the authors have considerable variation. Perhaps most trustworthy are Redmond's figures based on RPI and the effective-exchange-rate concept. These estimates suggest substantial overvaluation, which is consistent with the U.K. post-return experience of balance-of-payments deficits, deflation, and unemployment. Keynes predicted this in 1925, and he was right![14].

3.8.2 France

France re-adopted the gold standard on June 25, 1928, with a par value of 124.21 francs per pound sterling. This emanated from a gold par of exchange only slightly greater than one-fifth the prewar value, when mint parity was 25.225 francs per pound. The genesis of the new par value occurred in 1926, when several French officials made PPP computations yielding ranges of a stabilized rate. The best source of this history is Mouré (1996). In August, Pierre Quesnay's calculations, using WPI and Germany as the base country, yielded appropriate stabilization of 160–170 francs per pound. In November, Jacques Rueff's PPP computations employed both WPI and RPI price indexes, again with Germany as base country. He found the desired stabilization rate to be 120–145. In the same month, Charles Rist recommended the range 140–160.

In fact, the franc was appreciating in the foreign-exchange market. To stem this appreciation, at least temporarily, on December 20, 1926, Prime Minister Raymond Poincaré authorized the Bank of France to stabilize the rate via exchange-market intervention. This was a decision based on fear that appreciation would result in recession and unemployment and reduce Poincaré's political support within a coalition government. "PPP calculations did not decide the stabilization in December 1926" (Mouré, 1996, p. 144). However, as Mouré further comments, 'the economists' arguments were not without effect." Stabilization was at about 122 francs per pound and the return to the gold standard in 1928 at 124.21. These figures are close to the lower bound of Rueff's PPP computations. Mouré (1996, p. 148) writes: "With regard to choosing a rate of stabilization, PPP calculations offered evidence that was of interest but not decisive…"

Keynes (1930) [originally published in 1928], Cassel (1936), Walter (1951), and Sicsic (1992) provided estimates of undervaluation of the franc with respect to the British pound. The extent, not the direction, of deviation from PPP was the only issue. There is no doubt that undervaluation of the franc worsened the situation of the British, who overvalued the pound. At first, Keynes (1930, p. 114) judged that "the franc…fixed…at about one-fifth of its pre-war gold value…The figure finally chosen seems about right." Yet Keynes (1930, pp. 114–115) went on to state that a PPP computation would involve "a gold value of the franc nearer to one quarter (100 francs to the £) than to one-fifth of the pre-war value." This suggests about a 20% undervaluation (although Keynes did not state which price index he was applying). However, Keynes provided reasons—crudeness of French price indexes, room for domestic prices to rise, effect on export industry, budgetary implications, and avoidance of capital loss on foreign-exchange reserves of the Bank of France—why the French authorities were wise not to follow his computed PPP.

Using the WPI, Cassel, Walter, and Sicsic provided estimates of the undervaluation of the franc in the 6–12% range, the figure depending on the currency of comparison and the price index. These estimates are substantially below Keynes' figure of 20%.[15] Only Sicsic's RPI estimate, 28%, based on a nine-country effective exchange rate, exceeded that of Keynes.

It is not clear whether the French authorities *deliberately* undervalued the franc. According to Mouré, the concern was domestic macroeconomic stability, which explains why (i) the de facto stabilization rate in 1926 was undertaken to keep the franc from appreciating further, and (ii) the de jure stabilization rate in 1928 was close to the de facto rate established in 1926.

3.9 Conclusions

Why has PPP endured through the centuries and under the rubric of various and varying exchange-rate experiences and monetary standards? This survey of the application of PPP to historical experiences illustrates the controversial nature of PPP—and that nature is one reason for the durability and endurance of PPP. The second reason is that PPP is fundamentally a simple and intuitively appealing theory. The third reason is that it has an inherent concreteness that other exchange-rate theories lack.

Clearly, this survey shows that there is mixed empirical evidence for the applicability of PPP, whether in explaining exchange-rate behavior or in establishing new exchange-rate levels. Yet that result gives rise to a fourth reason for the robustness of PPP: *Whether or not PPP is deemed to hold* *empirically*, it is useful to know the extent to which the theory is valid. Measurement of *deviations* from PPP is important both for macroeconomic historians and for economic policy-makers. As Houthakker (1962, pp. 296–297) wrote: "All in all, it would be most unwise to ignore the unique insight which PPP calculations can afford."

Notes

1. However, the essay discusses only relative PPP, a reflection of "the PPP literature with a historical bent," itself a result of a paucity of data on absolute price levels in the pre-1940 period.
2. See Burns (1927, p. 417), Einzig (1970, p. 44), and Officer (1982, p. 27).
3. Citing the Roman historian Pliny, Einzig (1970, p. 45) reports that "merchants importing Indian goods sold them in Rome at a hundred-fold of what they had paid for them…the margin between the price of luxuries in their countries of origin and in their countries of destination was [some]thing like 10,000 per cent."
4. See the references in Officer, (1982, p. 194, n. 22).
5. This was pointed out originally by Keynes (1930, pp. 72–74; 1931, pp. 249–250 [originally published in 1925]).
6. Ahking (1990), Angell (1926), Cassel (1925a), Crump (1925), Farag and Ott (1964), Grilli and Kaminsky (1991), Hodgson (1972), MacDonald (1985a), Myhrman (1976), Stolper (1948), Taylor (1992), and Michael et al. (1996).
7. Aliber (1970), Angell (1926), Dulles (1929), Farag and Ott (1964), Myhrman (1976), Pippenger (1973), Rogers (1929), Wasserman (1936), and Sicsic (1992).
8. For Sweden, Anonymous (1921), Flux (1924), and Cassel (1925a, b); for Switzerland, Junge (1984) and Bleaney (1998); for Greece, Phylaktis (1990, 1992) and Georgoutsos and Kouretas (1992).
9. The relevant papers are those of Bachi (1925), Jacobson (1925a), Jacobson and Jaeger (1925), Wight (1925a, 1925b, 1925c), Wood (1925b, 1925c, 1925d1925a), Young (1925b, 1925c, 1925d, 1925e, 1925f, 1925g), and Jacobson et al. (1925).
10. The first group consists of Flux (1924), Furniss (1922), Graham (1930), Gregory (1925), Keynes (1923), Lester (1939), Robertson (1922), Tsiang (1959), U.S. Tariff Commission (1922), and Aliber (1962); the second, Hodgson and Phelps (1975), Hakkio (1984), Krugman (1978), Rogalski and Vinso (1977), Thomas (1973a, 1973b), and De Grauwe et al. (1985).
11. Bachi (1925), Copland (1930), Flux (1924), Gregory (1925), Katzenellenbaum (1925), and Thomas (1972).

12. Studies of the first group are Edison (1985), Frenkel (1978), Georgoutsos and Kouretas (2000), MacDonald (1985b); studies of the second group are Ardeni and Lubian (1989), Frenkel (1980), Michael et al. (1997), Taylor and McMahon (1988).
13. Excluded from the discussion are computations based on wages or export price indexes, as well as estimates emanating from more-general models of exchange-rate determination.
14. In February 1934, Czechoslovakia devalued the crown by 16 2/3%, with the devaluation rate based on a WPI PPP computation. Haberler (1961, p. 49, n. 37) comments that "exactly the same mistake was made [as in the United Kingdom in 1925]." The interpretation of Nurkse (1944, p. 128) was that the rate left no margin for economic expansion, putting downward pressure on the exchange value of the domestic currency. In any event, Czechoslovakia had to devalue a second time, in October 1936. In contrast, Belgium successfully devalued its franc in 1935. The devaluation rate of 28% was decided on the basis of PPP computations, with RPI as the decisive price concept. Further discussion of the Czech experience is in League of Nations (1936, pp. 49–52). The Belgian experience is discussed in League of Nations (1936, pp. 49–50), Nurkse (1944, p. 128), Garnsey (1945), Triffin (1937), and Officer (1982, pp. 143–144).
15. The estimates for Walter were computed by this author from Walter's data.

References

Ahking FW. Further results on long-run purchasing power parity in the 1920s. Eur Econ Rev 1990;34:913–919.

Aliber RZ. Speculation in the foreign exchanges: the European experience, 1919-1926. Yale Econ Essays 1962;2:171–245.

———. Speculation in the flexible exchange revisited. Kyklos 1970;23:303–313.

Angell JW. The theory of international prices: history, criticism and restatement. Cambridge (MA): Harvard University Press; 1926.

Anonymous. Is the Swedish exchange overvalued abroad? Skandinaviska Kreditaktiebolaget Quarterly Report; 1921. pp. 36–37.

Ardeni PG, Lubian D. Purchasing power parity during the 1920s. Econ Lett 1989;30:357–362.

Bachi R. Position of the currency and currency reform in Italy. Translated by Jacobson ML. In Young (1925a), Volume 1; 1925. pp. 161–196.

Bernholz P. Flexible exchange rates in historical perspective. Princeton (NJ): International Finance Section, Princeton University; 1982.

———. Monetary regimes and inflation: history, economic and political relationships. Cheltenham: Edward Elgar; 2003.

———, Manfred G, Heri EW. Historical experiences with flexible exchange rates: a simulation of common qualitative characteristics. J Int Econ 1985;19:21–45.
Bleaney M. Purchasing power parity in the 1920s: evidence for the Swiss franc. Appl Econ Lett 1998;5:239–241.
Bordo MD, Marcotte IA. Purchasing power parity in colonial America: some evidence for South Carolina 1732–1774: a comment on the Michener paper. Empirical Studies of Velocity, Real Exchange Rates, Unemployment and Productivity: Carnegie-Rochester Conf. Ser. Public Pol. 1987;27:311–324.
Bresciani-Turroni C. The economics of inflation: a study of currency depreciation in post-war Germany. London: George Allen and Unwin; 1937.
Broadberry SN. Purchasing power parity and the pound-dollar rate in the 1930s. Economica 1987;54:69–78.
——— Taylor MP. In: Broadberry SN, editor. Britain in the international economy. Cambridge: Cambridge University Press; 1992. pp. 369–395.
Bunting FH. The purchasing power parity theory reexamined. South Econ J 1939;5:282–301.
Burns AR. Money and monetary policy in early times. London: K. Paul, Trench, Trubner; 1927.
Cassel G. The present situation of the foreign exchanges. Econ J 1916;26:62–65.
———. Abnormal deviations in international exchanges. Econ J 1918;28:413–415.
———. The depreciation of the German mark. Econ J 1919;29:492–496.
——— Rates of exchange and purchasing-power parity, Skandinaviska Kreditaktiebolaget Quarterly Report; 1925a April. pp. 17–21.
——— Some recent experiences of monetary policy, Skandinaviska Kreditaktiebolaget Quarterly Report; 1925b Oct. pp. 55–58.
———. Exchanges, foreign. Supplementary Volume 1, Encyclopaedia Britannica. 13th ed. London: Encyclopaedia Britannica Company; 1926. pp. 1086–1089.
———. Post-war monetary stabilization. New York: Columbia University Press; 1928.
———. The downfall of the gold standard. Oxford: Clarendon Press; 1936.
Catão LAV, Solomou SN. Effective exchange rates and the classical gold standard adjustment. Am Econ Rev 2005;95:1259–1275.
Choudhry T, Luintel KB. The long-run behaviour of the real exchange rate: evidence from colonial Pennsylvania. Econ Lett 2001;74:25–30.
Copland DB. Credit and currency control. Melbourne: Melbourne University Press; 1930.
Crump N. The New York exchange: a reply to Professor Cassel. J R Stat Soc 1925;88:428–432.
De Grauwe P, Janssens M, Leliaert H. Real-exchange-rate variability from 1920 to 1926 and 1973 to 1982. Princeton (NJ): International Finance Section, Princeton University; 1985.

Delaplane WH. The Spanish Peseta Since 1913. [PhD dissertation]. Duke University; 1934.

Dimsdale NH. British monetary policy and the exchange rate 1920–1938. Oxf Econ Pap 1981;33:306–349.

Dulles EL. The French Franc, 1914–1928. New York: Macmillan; 1929.

Eagly RV. The Swedish and English bullionist controversies: a comparative study of the impact of events on ideas. In: Eagly RV, editor. Events, ideology and economic theory. Detroit: Wayne State University Press; 1968. pp. 13–31.

———. The Swedish bullionist controversy. Philadelphia (PA): American Philosophical Society; 1971.

Edison HJ. Purchasing power parity: a quantitative reassessment of the 1920s experience. J Int Money Finance 1985;4:361–372.

———. Purchasing power parity in the long run: a test of the Dollar/Pound exchange rate (1890–1978). J Money Credit Bank 1987;19:376–387.

Einzig P. The history of foreign exchange. 2nd ed. London: Macmillan; 1970.

Enders W. Unit roots and the real exchange rate before World War I: the case of Britain and the USA. J Int Money Finance 1989;8:59–73.

Farag AA, Ott DJ. Exchange rate determination under fluctuating exchange rates: some empirical evidence. In: Murphy JC, editor. Money in the international order. Dallas (TX): Southern Methodist University Press; 1964. pp. 84–105.

Flux AW. The foreign exchanges. London: P. S. King; 1924.

Frenkel JA. A monetary approach to the exchange rate: doctrinal aspects and empirical evidence. Scand J Econ 1976;78:200–224.

———. Purchasing power parity: doctrinal perspective and evidence from the 1920s. J Int Econ 1978;8:169–191.

———. Exchange rates, prices, and money: lessons from the 1920s. Am Econ Rev 1980;70:235–242.

Friedman M, Schwartz AJ. A monetary history of the United States, 1867–1960. Princeton (NJ): Princeton University Press; 1963.

Furniss ES. Foreign exchange: the financing mechanism of international commerce. Boston (MA): Houghton Mifflin; 1922.

Garnsey ME. Postwar exchange-rate parities. QJ Econ 1945;60:113–135.

Georgoutsos DA, Kouretas GP. Long-run purchasing power parity in the 1920s: the Greek experience. Appl Econ 1992;1301–1306.

———, Kouretas GP. The pound sterling and the Franc Poincaré in the1920s: long-run relationships, speculation and temporal stability. Appl Financ Econ 2000;10:471–482.

Graham FD. International trade under depreciated paper: the United States, 1862–79. QJ Econ 1922;36:220–273.

———. Exchange, prices, and production in hyper-inflation Germany, 1920–1923. Princeton (NJ): Princeton University Press; 1930.

———. Recent movements in international price levels and the doctrine of purchasing power parity. J Am Stat Assoc 1935;30:159–166.
Gregory TE. Foreign exchange before, during, and after the War. London: Oxford University Press; 1925.
———. The first year of the gold standard. London: Ernest Benn; 1926.
Grilli V, Kaminsky G. Nominal exchange rate regimes and the real exchange rate: evidence from the United States and Great Britain, 1885–1986. J Monet Econ 1991;27:191–212.
Grubb FW. Creating the U.S. Dollar currency union, 1748–1811: a quest for monetary stability or a usurpation of state sovereignty for personal gain. Am Econ Rev 2003;93:1778–1798.
———. State 'Currencies' and the transition to the U.S. Dollar: reply—including a new view from Canada. Am Econ Rev 2005;95:1341–1348.
———. Testing for the economic impact of the U.S. constitution: purchasing power parity across the colonies versus across the states, 1748–1811. J Econ Hist 2010;70:118–145.
Haberler G. The theory of international trade. London: William Hodge; 1936.
———. A survey of international trade theory. Princeton (NJ): International Finance Section, Princeton University; 1961.
Hakkio CS. A re-examination of purchasing power parity: a multi-country and multi- period study. J Int Econ 1984;17:265–277.
Hasan MS. Univariate time series behaviour of the real exchange rate: evidence from colonial India. Econ Lett 2004;84:75–80.
Heckscher EF. Sweden. In: Heckscher EF, editor. Sweden, Norway, Denmark and Iceland in the World War. New Haven (CT): Yale University Press; 1930. pp. 125–217.
Hegwood ND, Papell DH. Purchasing power parity under the gold standard. South Econ J 2002;69:72–91.
Hodgson JS. An analysis of floating exchange rates: the dollar-sterling rate, 1919–1925. South Econ J 1972;39:249–257.
———, Phelps P. The distributed impact of price-level variation on floating exchange rates. Rev Econ Stat 1975;57:58–64.
Houthakker HS. Exchange rate adjustment. In: Factors Affecting the United States Balance of Payments, Compilation of studies prepared for the Subcommittee on International Exchange and Payments, Joint Economic Committee, 87th Congress. 2nd session. Washington (DC): U.S. Government Printing Office; 1962. pp. 287–304.
Jacobson ML. Russia, In Young (1925a), Volume 2; pp. 185–203.
———, Jaeger R. Germany, In Young (1925), Volume 1; 1925. pp. 387–343.
———, Jaeger R, Young JP. Italy, In Young (1925a), Volume 1; 1925. pp. 347–386.

Junge G. Purchasing power parity in the 1920s and the 1970s. Eur Econ Rev 1984;26:73–82.

Katano H. Econometric determination of foreign exchange rate of Japan for 1926-1935. Kobe Econ Bus Rev 1956;3:9–25.

———. Reconstruction of the theory of purchasing power parity. Kobe Econ Bus Rev 1957;4:125–148.

Katzenellenbaum SS. Russian currency and banking, 1914–1924. London: P.S. King; 1925.

Keynes JM. Official figures relating to inflation. Econ J 1919;29:504–506.

———. A tract on monetary reform. London: Macmillan; 1923.

———. A treatise on money, Volume 1. London: Macmillan; 1930.

———. Essays in persuasion. London: Macmillan; 1931.

Kindahl JK. Economic factors in specie resumption: the United States, 1865–79. J Pol Econ 1961;69:30–48.

Krugman PR. Purchasing power parity and exchange rates: another look at the evidence. J Int Econ 1978;8:397–407.

League of Nations. Monetary Review, Money and Banking, 1935/36, Volume 1, Geneva; 1936.

Lester RA. Monetary experiments: early American and recent Scandinavian. Princeton (NJ): Princeton University Press; 1939.

MacDonald R. The Norman Conquest of $4.86 and the asset approach to the exchange rate. J Int Money Finance 1985a;4:373–387.

———. Are deviations from purchasing power parity efficient? Some further answers. Weltwirtsch Arch 1985b;121:638–645.

Matthews KGP. Was sterling overvalued in 1925? Econ Hist Rev 1986;39:572–587.

McCloskey DN, Zecher JR. The success of purchasing power parity: historical evidence and its implications for macroeconomics. In: Bordo MD, Schwartz AJ, editors. A retrospective on the classical gold standard, 1821–1931. Chicago: University of Chicago Press; 1984. pp. 121–150.

Michael P, Nobay AR, Peel DA. Exchange rate dynamics and monetary reforms: theory and evidence from Britain's return to gold. In: Bayoumi T, Eichengreen B, Taylor MP, editors. Modern perspectives on the gold standard. Cambridge: Cambridge University Press; 1996. pp. 341–362.

———. Nobay AR, Peel DA. Transactions costs and nonlinear adjustment in real exchange rates: an empirical investigation. J Polit Econ 1997;105:862–879.

Moggridge DE. The return to gold, 1925. Cambridge: Cambridge University Press; 1969.

———. British monetary policy 1924–1931: the Norman conquest of $4.86. Cambridge: Cambridge University Press; 1972.

Mouré K. Undervaluing the Franc Poincaré. Econ Hist Rev 1996;49:137–153.

Myhrman J. Experiences of flexible exchange rates in earlier periods: theories, evidence and a new view. Scand J Econ 1976;78:169–196.

Nachane DM, Hatekar NR. The bullionist controversy: an empirical reappraisal. Manch Sch 1995;43:412–425.

Nurkse R. International currency experience, League of Nations. Princeton (NJ): Princeton University Press; 1944.

Officer LH. The floating dollar in the greenback period: a test of theories of exchange-rate determination. J Econ Hist 1981;629–650.

———. Purchasing power parity and exchange rates: theory, evidence and relevance. Greenwich (CT): JAI; 1982.

———. The bullionist controversy: a time-series analysis. Int J Finance Econ 2000;5:197–209.

———. Purchasing power parity: from ancient times to World War II. In EH.Net Encyclopedia, Economic History Services, EH.Net, 2006.

Phylaktis K. Purchasing power parity: the Greek evidence from the 1920s. Greek Econ Rev 1990;12:148–156.

———. Purchasing power parity and cointegration: the Greek evidence from the 1920s. J Int Money Finance 1992;11:502–513.

Pippenger JE. Speculation in the flexible exchange re-revisited. Kyklos 1973;26:613–618.

Redmond J. The sterling overvaluation in 1925: a multilateral approach. Econ Hist Rev 1984;37:520–532.

Robertson DH. Money. New York: Harcourt, Brace; 1922.

Rogalski RJ, Vinso JD. Price level variations as predictors of flexible exchange rates. J Int Bus Stud 1977;8:71–81.

Rogers JH. The process of inflation in France, 1914-1927. New York: Columbia University Press; 1929.

Sabaté M, Gadea MD, Serrano JM. PPP and structural breaks: the peseta-sterling rate, 50 years of a floating regime. J Int Money Finance 2003;22:613–627.

Sayers RS. The return to gold, 1925. In: Pressnell LS, editor. Studies in the industrial revolution. London: Athlone Press; 1960. pp. 313–327.

Schweigert TE. Nominal and real exchange rates and purchasing power parity during the Guatemalan float, 1897-1922. J Econ Dev 2002;27:127–142.

Sicsic P. Was the franc Poincaré deliberately undervalued? Explor Econ Hist 1992;29:69–92.

Sort MS, Gadea MD, Serrano JM. The Spanish peseta versus the pound sterling, the French franc and the US dollar (1870–1935): a long floating experience. Appl Financ Econ Lett 2005;1:95–99.

Stolper WF. Purchasing power parity and the pound sterling from 1919–1925. Kyklos 1948;2:240–269.

Taylor MP. Dollar-sterling exchange rate in the 1920s: purchasing power parity and the Norman conquest of $4.86. Appl Econ 1992;24:803–811.

———, McMahon PC. Long-run purchasing power parity in the 1920s. Eur Econ Rev 1988;32:179–197.
Thomas LB. Some evidence on international currency experience, 1919–1925. Nebr J Econ Bus 1972;11:145–155.
———. Speculation in the flexible exchange revisited—another view. Kyklos 1973a;26:143–150.
———. Behavior of flexible exchange rates: additional tests from the post-World War I episode. South Econ J 1973b;40:167–182.
Thompson GR. Expectations and the greenback rate, 1862–1878. [PhD dissertation]; University of South Florida; 1972.
Triffin R. La théorie de la surévaluation monétaire et la devaluation belge. Bull d'Institut de Rech Econ Soc 1937;9:19–52.
Tsiang S-C. Fluctuating exchange rates in countries with relatively stable economies: some European experiences after World War I. Int Monet Fund Staff Pap 1959;7:244–273.
United States Tariff Commission. Depreciated exchange and international trade. Washington (DC): Government Printing Office; 1922.
Walter JT. Foreign exchange equilibrium. Pittsburgh (PA): University of Pittsburgh Press; 1951.
Wasserman MJ. The compression of French wholesale prices during inflation, 1919–1926. Am Econ Rev 1936;26:62–73.
White HG. Foreign exchange rates and internal prices under inconvertible paper currencies. Am Econ Rev 1935;25:259–272.
Whitaker JK, Hudgins MW Jr. The floating pound sterling of the nineteen-thirties: an econometric study. South Econ J 1977;43:1478–1485.
Wight EVD. Austria, In Young (1925a), Volume 2; 1925a. pp. 9–23.
———. Belgium, In Young (1925a), Volume 2; 1925b. pp. 27–49.
———. Netherlands, In Young (1925a), Volume 2; 1925c. pp. 125–139.
Wood E. Denmark, In Young (1925a), Volume 2; 1925a. pp. 79–97.
———. Norway, In Young (1925a), Volume 2; 1925b. pp. 143–159.
———. Sweden, In Young (1925a); Volume 2; 1925c. pp. 245–266.
———. Switzerland, In Young (1925a), Volume 2; 1925d. pp. 267–289.
Yeager LB. Fluctuating exchange rates in the nineteenth century: the experiences of Austria and Russia. In: Mundell RA, Swoboda AK, editors. Monetary problems of the international economy. Chicago: University of Chicago Press; 1969. pp. 61–89.
——— International monetary relations: theory, history, and policy. 2nd ed. New York: Harper & Row; 1976.
Young JP. European currency and finance. Two volumes. Washington (DC): Government Printing Office; 1925a.
———. Czechoslovakia, In Young (1925a), Volume 2; 1925b. pp. 55–69.
———. France, In Young (1925a), Volume 1; 1925c. pp. 307–346.

———. Hungary, In Young (1925a), Volume 2; 1925d. pp. 103–124.
———. Poland, In Young (1925a), Volume 2; 1925e. pp. 163–184.
———. Spain, In Young (1925a), Volume 2; 1925f. pp. 223–239.
———. Great Britain, In Young (1925a), Volume 1; 1925g. pp. 273–306.
———. International trade and finance. New York: Ronald; 1938.

CHAPTER 4

Afterword to Part I

4.1 Salamancans and Gerard Malynes (Chapter 1)

A reason why the origin of purchasing power parity (PPP) as a theory is in the economic-history rubric is that both the Salamancans and Gerard Malynes were well aware of the tremendous inflow of gold and silver from the Americas, and of the resultant increase in coin stock.[1] While Spain was a "first-order receiver" of the precious metals and England a "second-order receiver" (using Nuno Palma's terminology), the upward trend in commodity prices over the sixteenth century is common to both countries.[2] Indeed, by one calculation (Douglas Fisher, 1989, p. 895), prices actually increased at a *higher* rate in England than Spain; but the well-known Spanish data (per Earl J. Hamilton, 1934) likely underestimate inflation.[3]

The Salamancans were late Scholastics. As such, they were professors and theologians and may be presumed to have led mundane—if not austere—personal lives.[4] The same cannot be said of Malynes. His colorful existence is summarized all too briefly in Chapter 1. Among his dubious business ventures were producing token farthings and transacting in the cargo of a captured Spanish ship. The latter activity resulted in Malynes found to be an embezzler, for which he was sentenced to jail, but "where he continued to carry on mercantile ventures by correspondence" (Lynn Muchmore, 1969, p. 338).[5] Yet he continued to enjoy government connections and continued his business activity.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_4

"The commercial environment which Malynes knew was dominated by men who thrived on the foggy intrigue and the speculative fortunes made or lost through the caprice of war…Malynes was fully practiced in the ways of the unscrupulous, and it was an appreciation of their power that he brought to his economic pamphlets" (Muchmore, 1969, p. 339). Further, Malynes was a prolific writer—of pamphlets, treatises, and petitions and memoranda submitted to the Privy Council (Muchmore, 1969, p. 337).

The argument that Malynes was more concerned with denouncing the "manipulations of exchange dealers" rather than developing "a complete formulation of a full cycle of the self-regulating mechanism" (Jacob Viner, 1937, p. 76) is an exaggeration and patently unfair to the later Malynes. However, "even after he ceased to attack the bankers in his writings, he relied upon his own impressions that the market was never impersonal and that prices were influenced by men who skillfully deployed financial power" (Muchmore, 1969, p. 342). In advocating the restoration and extension of exchange control to support officially fixed exchange rates at the mint parity, Malynes would minimize the influence of dealers in foreign exchange.

4.2 Gustav Cassel (Chapter 2)

I have profound admiration for Gustav Cassel's contribution to the development of PPP both as a theory and in empirical application. Indeed, he is one of the three scholars who are my intellectual heroes: Gustav Cassel, Gottfried Haberler, and Joseph Schumpeter. It follows that I am sensitive to unwarranted criticism of their work.

Michael Michaely (1982) argues that the *relative* PPP theory is the basis of Cassel's PPP framework. According to Michaely, Cassel developed absolute PPP only because, given Sweden's replacement of the gold standard during World War I with a floating exchange rate, mint parity is replaced by absolute PPP. Fundamentally, the price level of commodities substitutes for the gold price in the domestic and foreign countries.

On the contrary, in Officer (1982) I show that Cassel's own writings demonstrate that he developed PPP long before World War I, that his basic PPP theory was the *absolute* version, and that relative PPP was promulgated as an expedient, so that the theory could be tested and put to policy use. After all, in contrast to measures of inflation (price indexes), price-level statistics were virtually impossible to obtain. The former is all that is needed for relative PPP, the latter is required for absolute PPP.

However, on reflection, I may have been too harsh on Michael Michaely, who writes: “What had first led him [Gustav Cassel] to the development of the [PPP] theory was probably the fact that Sweden had formally abandoned (in early 1916) the gold standard—a fact repeatedly mentioned in Cassel’s papers. A substitute for the ‘gold parity’ must then be invented” (Michaely, 1982, p. 244). From the start of World War I, exchange-rate fluctuations of currencies of gold-standard countries combined with differing rates of inflation in these countries created an environment clearly conducive to the relative PPP theory.[6] So it is surprising indeed that Cassel had developed the PPP theory—and specifically the *absolute* PPP theory—almost a decade prior to the war.[7] Thus—contrary to Malynes and the Salamancans—there is an unusual contradiction between economic history and economic thought!

4.3 Other Studies (Chapter 3 Extended)

Overriding the rule in Chapter 3 that “a study’s time period must fully antedate the year 1940,” interesting work of a few other authors warrants discussion. There is the amazing finding of Hendrik S. Houthakker (1962, p. 297), who used PPP for a precise estimate of the mark/dollar equilibrium exchange rate. This was 1962, when the Bretton Woods system was functioning, and Germany had revalued its currency from 4.2 to 4 marks per dollar a year earlier. Working with absolute-PPP cost-of-living data from the German Statistical Office, Houthakker finds that the equilibrium rate in 1962 is 3.112 DM—a dollar overvaluation of 22.2 percent. Then what happened? Following episodes of managed float and revaluation, the mark reached an average level of 3.108 over 1970–1974.[8] Houthakker’s PPP estimate is amazingly accurate—identical to the later relatively free exchange rate to two decimal places. A plus to the PPP theory over the medium term!

James R. Lothian and Mark P. Taylor (1996, 2008) initiated and represent a body of literature that tests PPP via sophisticated time-series analysis applied to the long-run dollar-sterling exchange rate, with annual data sets 1791–1990 and 1820–2001.[9] They cite Milton Friedman and Anna Jacobson Schwartz (1963, pp. 678–679): “the stability of basic economic relations” is reflected in “the behavior [limited range] of relative prices in the United States and Great Britain adjusted for changes in the exchange rate between the dollar and the pound…from 1871 to 1949.“ Indeed, Lothian and Taylor (1996, p. 505) find that PPP is "a useful empirical first approximation," the more so when relative productivities are incorporated (Lothian and Taylor, 2008).

Notes

1. For the relationship between bullion imports and coin production, see Nuno Palma (2020, pp. 364–368). There is no evidence that the Salamacans (as a group) and Malynes were cognizant of each-other's writings; so it is likely that their contributions to the quantity theory and PPP were subjectively original.
2. See the graph in Douglas Fisher (1989, p. 894).
3. See Palma (2020, p. 373).
4. However, some Salamancans, including Azpilcueta de Navarro himself, were active in Church or government affairs. See Raymond de Roover (1974, p. 312).
5. For biographies of Malynes, one may consult William A. S. Hewins (1893) and Muchmore (1969, pp. 338–339).
6. "Already at the outbreak of the War the gold standard began to be abandoned, and the neutral currencies were reduced to free paper standards in which gold and such currencies as the pound and the dollar were quoted above par" (Cassel, 1922, p. 98). Appreciation of the Swedish krona against the British pound from the start of World War I is specifically mentioned.
7. For assessments of Cassel's contributions to PPP and economics in general, see Hans Brems (1989), Denis V. Kradochnikov (2013), and Chapter 2.
8. Susan B. Carter and others (2006, series Ee626). The period 1970–1974 is selected, because (a) 1970 is the first year of the exchange rate below 3.9, implying greater freedom (less management) of the floating mark, and (b) five years is a sufficiently long interval for an equilibrium exchange rate to be discerned.
9. The franc-sterling exchange rate is also examined.

References

Bernholz, Peter (2015). *Monetary Regimes and Inflation*, 2nd ed. Northampton, MA: Edward Elgar.

Brems, Hans (1989). "Gustav Cassel Revisited." *History of Political Economy* 21 (June), pp. 165–178.

Carter, Susan B., and others (2006). *Historical Statistics of the United States, Millennial Edition*. Cambridge: Cambridge University Press.

Cassel, Gustav (1922). *Money and Foreign Exchange After 1914*. London: Constable.

De Roover, Raymond (1974). *Business, Banking and Economic Thought*, Julius Kirshner, ed. Chicago: University of Chicago Press.

Fisher, Douglas (1989). "The Price Revolution: A Monetary Interpretation." *Journal of Economic History* 49 (December), pp. 883–902.

Friedman, Milton, and Anna Jacobson Schwartz (1963). *A Monetary History of the United States 1867–1960*. Princeton: Princeton University Press.

Hamilton, Earl J. (1934). *American Treasure and the Price Revolution in Spain, 1501–1650*. Cambridge, MA: Harvard University Press.

Hewins, William A. S. (1893). "Malynes, Malines, or De Malines, Gerard." In Stephen Leslie, ed., *Dictionary of National Biography*, vol. 36. London: Smith, Elder.

Houthakker, H. S. (1962). "Exchange Rate Adjustment." In Joint Economic Committee, *Factors Affecting the United States Balance of Payments*, pp. 287–304. Washington: Government Printing Office.

Kadochnikov, Denis V. (2013). "Gustav Cassel's Purchasing Power Parity Doctrine in the Context of his Views on International Economic Policy Coordination." *European Journal of the History of Economic Thought* 20 (December), pp. 1101–1121.

Lothian, James R., and Mark P. Taylor (1996). "Real Exchange-Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries." *Journal of Political Economy* 104 (June), pp. 488–509.

——— and Mark P. Taylor (2008). "Real Exchange Rates Over the Past Two Centuries: How Important is the Harrod-Balassa-Samuelson Effect?" *Economic Journal* 118 (October), pp. 1742–1763.

Michaely, Michael (1982). "Gustav Cassel's Early Purchasing-Power-Parity Theory: A Note." *History of Political Economy* 14 (Summer), pp. 242–245.

Muchmore, Lynn (1969). "Gerrard de Malynes and Mercantile Economics." *History of Political Economy* 1 (Fall), pp. 336–358.

Officer, Lawrence H. (1982). "The Relationship Between the Absolute and the Relative PPP Theory of Cassel." *History of Political Economy* 14 (Summer), pp. 251–255.

Palma, Nuno (2020). "American Precious Metals and Their Consequences for Early Modern Europe." In Stefano Battilossi, Youssef Cassis and Kazuhiko Yago, eds. *Handbook of the History of Money and Currency*, pp. 363–382. Singapore: Springer Nature.

Viner, Jacob (1937). *Studies in the Theory of International Trade*. New York: Harper.

PART II

Purchasing Power Parity: Empirical Studies

CHAPTER 5

Absolute and Relative Purchasing Power Parity

5.1 The Relationship Between Absolute and Relative Purchasing Power Parity

Republished with permission of MIT Press, from *Review of Economics and Statistics*, Vol. 60, No. 4 (November 1978): pp. 562–568; permission conveyed through Copyright Clearance Center, Inc.

The absolute purchasing power parity (PPP) theory asserts that the equilibrium exchange rate (number of units of domestic currency per unit of standard currency) is determined by the ratio of the price level of the domestic country to the price level of the standard country.[1] This ratio is itself called the absolute PPP. The relative PPP theory states that the ratio of the equilibrium exchange rate in a current period (t) to the equilibrium exchange rate in a base period (o) is determined by the ratio of the domestic country's price index in period t to the standard country's price index in period t, where both indexes are measured relative to period o.

Suppose that the absolute PPP theory is fulfilled in both periods t and o. Then the relative PPP theory may be restated as follows: the ratio of absolute PPP in period t to absolute PPP in period o is determined by the ratio of the domestic country's price index to the standard country's price index, where both indexes are measured in period t relative to period o. The question immediately arises, however, whether the relative PPP

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_5

theory has now become a truism. Is the current/base-period absolute-PPP ratio *identically equal* to the domestic/standard-country price-index ratio? It is the purpose of this article, first, to demonstrate that the restated relative PPP theory is *not* a truism and, second, to provide an empirical test of this interpretation of the PPP hypothesis.

The restatement of the relative PPP theory separates its validity from that of the absolute PPP theory. Relative PPP becomes concerned only with the *movement* from one *potential* exchange-rate equilibrium to another. Whether the exchange rate is actually in equilibrium (à la PPP) at the two end points of the time period becomes the purview of the absolute PPP hypothesis, and is an issue beyond the confines of this paper. But it must be shown that the restated PPP hypothesis is an operational theory rather than a truism.

5.1.1 Proof That the Restated PPP Theory Is Not a Truism

Consider the following notation:

P_j = absolute purchasing power parity in period j, number of units of domestic currency per unit of standard currency
L_j = price level of the domestic country in period j
$L_j{}^s$ = price level of the standard country in period j
I_j = price index of the domestic country in period j relative to period o
$I_j{}^s$ = price index of the standard country in period j relative to period o
p_{ij} = price of commodity i in the domestic country in period j
$P_{ij}{}^s$ = price of commodity i in the standard country in period j
w_{ij} = weight of the price of commodity i in the domestic country's price level in period j
$w_{ij}{}^s$ = weight of the price of commodity i in the standard country's price level in period j

Then $w_{io}\left(w_{io}^s\right)$ is the weight of the price of commodity i in the domestic (standard) country's price *index* for all time periods, in particular, for periods o and t.

By definition,

$$L_j \equiv \sum_i w_{ij} p_{ij} \qquad j = o, t$$

$$L_{j^s} \equiv \sum_i w_{ij^s} p_{ij^s} \qquad j = o, t$$

$$P_j \equiv L_j / L_{j^s} \qquad j = o, t$$

$$I_t \equiv \frac{\sum_i w_{io} p_{it}}{\sum_i w_{io} p_{io}}$$

$$I_{t^s} \equiv \frac{\sum_i w_{io^s} p_{it^s}}{\sum_i w_{io^s} p_{io^s}}$$

The restated PPP theory is a truism if and only if

$$\frac{P_t}{P_o} \equiv \frac{I_t}{I_{t^s}}, \tag{5.1}$$

i.e., if and only if

$$\frac{\sum_i w_{it} p_{it} / \sum_i w_{it^s} p_{it^s}}{\sum_i w_{io} p_{io} / \sum_i w_{io^s} p_{io^s}} \equiv \frac{\sum_i w_{io} p_{it} / \sum_i w_{io} p_{io}}{\sum_i w_{io^s} p_{it^s} / \sum_i w_{io^s} p_{io^s}}$$

or

$$\frac{\sum_i w_{it} p_{it}}{\sum_i w_{it^s} p_{it^s}} \equiv \frac{\sum_i w_{io} p_{it}}{\sum_i w_{io^s} p_{it^s}}$$

or

$$w_{it} = w_{io} \text{ and } w^s_{it} = w^s_{io} \text{ for all } i \tag{5.2}$$

What is the interpretation of (5.2)? It states that, for both the domestic and standard country, the weights of the component prices in the *price level* of the country are the same in the current as in the base period. For the country's price *index*, of course, the weighting pattern in the current period is, by definition, equal to that in the base period. But, for the price level, equal weighting patterns in the two periods would exist only in a very special case.

Since under the PPP theory a country's price level covers its entire output of commodities, the weights of the price level reflect the production pattern of the country. Assume the absence of money illusion, so that a mere change in units of measurement makes no difference to real economic behavior. Then the only situation in which production would be distributed among all commodities in an unvarying proportion in two time periods would be *the absence of any changes in relative prices* in the economy. Therefore (5.2) holds if and only if, for both the domestic and standard country, all individual prices in the country in period t are a constant multiple of the prices in period o, i.e.,

$$p_{it} = c \cdot p_{io} \text{ and } p_{it}^{s} = c^{s} \cdot p_{io}^{s} \text{ for all } i \tag{5.3}$$

where c and c^s are positive constants.

What has been demonstrated? Assuming the absence of money illusion, identity (5.1) is fulfilled if and only if Eqs. (5.3) hold, i.e., there is pure inflation or deflation in both the domestic and standard country. Furthermore, for (5.1) to hold, the only permissible real change in the economies is an equiproportional increase or decrease in the production of every commodity.

The condition for the restated PPP theory to be a truism is stringent indeed. In practice, prices of individual commodities do not move uniformly with the general price level and therefore the relative production of individual commodities (weights of the individual prices in the price level) also change. So Eq. (5.1) in its identity form cannot be expected to be fulfilled in the real world.

5.1.2 Alternative Price-Level Concepts of PPP

Let $P \equiv P_t/P_o$ and $D \equiv I_t/I_t^s$. It has been shown that the restated relative PPP theory is not an identity, that is, P is not identically equal to D. Therefore the theory may legitimately be tested empirically, and it has the general form $P = h(D)$, where h is an increasing function.

Two separate data sets are used to generate samples of observations on P and D.[2] The first data set involves a gross domestic product (GDP) price-level concept for the absolute-PPP computation (P) and, correspondingly, the GDP deflator as the price measure for constructing variable *D*. The United States is the standard country. The second data set employs a cost-of-living (COL) concept of PPP and the consumer price index as the price measure, with Germany as the standard country.

There are several reasons why the GDP-concept samples are deemed superior to the COL-concept samples for the purpose of testing the PPP theory. An empirical reason is that the United States, as the dominant country in the world economy, can be construed as the optimal standard country for any broad group of domestic countries. There also exist two theoretical arguments in favor of the GDP-concept data set, one on the consumption side, the other on the production side of the economy. To the extent that the PPP theory is justified by the existence of arbitrage and substitutability of commodities in consumption (broadly construed), the price concept underlying PPP should be as comprehensive as possible. Therefore a GDP measure, encompassing all output of the economy, is preferred to a COL measure, which restricts pricing to those commodities purchased by households. On the production side, it can be argued that a unit-factor-cost concept is the most appropriate methodology for absolute PPP (Houthakker 1962, pp. 293–294). Now, under certain assumptions, a unit-factor-cost concept of PPP is equivalent to a PPP based on price levels that are production-weighted averages of commodity prices in each country, implying a GDP price-level measure for PPP (Houthakker 1962, p. 296; Officer 1974, pp. 871–872; 1976, pp. 11–12). However, for equivalence with a COL concept of PPP, i.e., with the use of household consumption weights in the construction of price levels, additional—and more stringent—assumptions are required (Officer 1976, pp. 12–13).

On the other hand, the theoretical argument in Sect. 5.1.1 implies that the appropriate price index for the construction of variable *D* is base-weighted rather than current-weighted. Yet the only available GDP price index is the current-weighted deflator. In contrast, the consumer price index, used in association with the COL-concept PPP, is base-weighted, although there may be changes in the weighting pattern at discrete points in time.

While, on balance, the GDP concept may be construed as the preferred foundation for PPP computation, data availability limits the size of the samples that can be generated on this basis. Observations on *P* and *D* are collected for 8 countries in the 1950–1955 period (that is, with 1950 as the base year and 1955 as the current period), 4 countries in the 1967–1970 period, and 4 in the 1950–1970 period. Thus there are three distinct samples, and the countries composing each sample are listed in the first column of Table 5.1.

Use of the COL concept enables the assembling of a much larger data set. Unfortunately, there is no uniformity in base and current periods. So samples are delineated on the basis of the *duration* between base and current period: (i) less than 10 years, a 15-observation sample, (ii) 10 to 19 years, 9 observations, and (iii) 20 years or more, 6 observations. For each sample, the observations are identified by country, base period, and current period in the first column of Table 5.2.

Table 5.1 Errors of strong PPP and naive models: GDP concept

Country	*Percentage error*	
	$\frac{P-D}{P}$	$\frac{P-1}{P}$
1950–1955 period		
Belgium	3.16	0.73
Denmark	−1.15	5.76
France	3.57	21.43
Germany	−6.35	−2.68
Italy	−3.32	3.54
Netherlands	0	7.54
Norway	−2.67	13.19
United Kingdom	−2.52	8.63
Average of Absolute Values	2.84	7.94
1967–1970 period		
Hungary	−3.68	−7.36
Japan	2.03	3.66
Kenya	−1.10	−7.97
United Kingdom	−2.01	1.01
Average of Absolute Values	2.21	5.00
1950–1970 period		
France[a]	−0.44	41.46
Germany[a]	−7.52	3.76
Italy[a]	−16.20	5.83
United Kingdom[a]	−14.77	14.77
Average of Absolute Values	9.73	16.46

[a]Observation excluded from maximum-size sample

Table 5.2 Errors of strong PPP and naive models: COL concept

Country	*Percentage error*	
	$\frac{P-D}{P}$	$\frac{P-1}{P}$
Computational period: less than 10 years		
Austria (1954–1960)	1.35	2.13
Austria (1960–1968)	9.01	15.55
France (1952–1958)	7.17	16.33
Israel (1957–1961)	25.38	32.34
Israel (1961–1969)	4.23	22.87
Italy (1967–1972)	−9.17	−7.35
Netherlands (1953–1960)	−11.79	−2.97
Netherlands (1960–1967)	4.93	10.98
New Zealand (1956–1965)	7.04	10.55
Norway (1954–1960)	0.31	4.11
Sweden (1952–1959)	−16.52	−2.67
Switzerland (1952–1957)	5.00	4.82
Switzerland (1957–1964)	−7.70	−7.55
United Kingdom (1953–1961)	9.30	15.02
Soviet Union (1954–1958)	−15.29	−22.73
Average of Absolute Values	8.95	11.86
Computational period: 10 to 19 years		
Austria (1954–1968)[a]	10.24	17.36
Denmark (1958–1975)	12.51	44.97
France (1958–1972)	7.83	26.10
Israel (1957–1969)[a]	28.54	47.81
Italy (1952–1967)	−5.36	12.02
Netherlands (1953–1967)[a]	−6.28	8.33
Norway (1960–1974)	8.02	27.36
Switzerland (1964–1974/75)	−8.14	4.12
United Kingdom (1961–1975)	12.57	43.71
Average of Absolute Values	11.05	25.75
Computational period: 20 years or more		
France (1952–1972)[a]	14.44	38.17
Italy (1952–1972)[a]	−15.02	5.55
Norway (1954–1974)[a]	8.31	30.35
Switzerland (1952–1974/75)[a]	−10.63	1.85
United Kingdom (1953–1975)[a]	20.70	52.17
United States (1953–1973)	−2.31	−5.47
Average of Absolute Values	11.90	22.26

[a]Observation excluded from maximum-size sample

5.1.3 Empirical Analysis of the Strong PPP and Naive Models

The "strong PPP model" is obtained by the inclusion of an error term in Eq. (5.1):

$$P = D + \epsilon_1 \tag{5.4}$$

In general, an error term is denoted by a subscripted ϵ. The strong PPP model is tested against a corresponding naive model, called the "strong naive model":

$$P = 1 + \epsilon_2 \tag{5.5}$$

The strong naive model, in effect, predicts P_t, the absolute PPP in period t, by P_0, the base-period PPP. Price indexes in the domestic and standard countries are assigned no role in predicting absolute PPP in the current period relative to the base period.

A comparison of the performance of the strong PPP and strong naive models can be made by calculating their percentage errors in ex post prediction. These errors are $P - D/P$ and $P - 1/P$, respectively. They are shown in Tables 5.1 and 5.2 for the samples based on the GDP and COL concepts, respectively. A positive error implies an underestimate of P, while a negative error implies an overestimate. Then under the GDP concept, the PPP model has a tendency to overestimation, while the naive model has the opposite tendency. With a total of 16 observations over the three samples, the PPP model overpredicts P in 12 cases, with the naive model overpredicting in only 3 cases. The implication is that the relative price level between a domestic and a standard country in a current compared to a base period, tends to be less than that indicated by the corresponding ratio of price indexes between the countries. The relative PPP hypothesis tends to predict too great a change in absolute PPP on the basis of changes in price indexes.

This result does not carry over to the COL-concept samples. Except for the 20-years-or-more computational period, the PPP model *underestimates* P at double the rate that it overestimates it, while the naive model tends to underpredict P in all samples, and with greater overall frequency. There is no apparent reason for the expected direction of the forecast error to vary with the price-level concept of PPP.

In any event, the direction of a prediction error is less relevant than the amount of the error, especially for comparison with a naive model.

Consider the GDP-concept samples first. In terms of *absolute* percentage errors, the PPP model is superior to the naive model (i.e., has a lower percentage error) for 10 of the 16 observations over all samples, it is inferior to the naive model for 5 observations, and there is an equal percentage error for 1 observation. For both shorter time periods (1950–1955 and 1967–1970), the PPP model is superior by a three-to-one margin, under this ordinal criterion. Only for the longer time period (1950–1970) does the naive model outperform the PPP model (by a two-to-one margin, with one tie).

The ordinal superiority of the PPP model is stronger for the COL concept. Over all samples, the PPP model has a lower absolute percentage error than the alternative model in 22 of 30 observations. The PPP model is superior in 8 of 9 observations for the 10-to-19-years sample, and by a two-to-one margin in the other samples.

A cardinal measure of performance of the models is the *average* of the absolute values of the percentage errors. For each sample, this average is shown in Tables 5.1 and 5.2. Now the PPP model is superior to the naive model in all samples for both data sets. Especially noticeable in the GDP-concept samples are the low average percentage errors of the PPP model for the shorter time periods, with the average between 2 and 3%. This result is particularly impressive when coupled with the fact that the highest absolute error is below 7% (and the second highest below 4%) for the 1950–1955 period, and below 4% for the 1967–1970 period.

In contrast, an average absolute error of about 9 3/4% marks a less prominent performance of the PPP model for the 1950–1970 period, even though this result is superior to that of the naive model, which has an average error of nearly 16 1/2%. The inferior result for the 1950–1970 period is not unexpected, because with a longer time period there is greater scope for changes in the price-quantity structure underlying a country's price level, thus reducing the applicability of the PPP model.

Turning to the COL-concept samples, the average absolute error of the PPP model is approximately 9%, 11%, and 12%, respectively, for the three samples in order of duration of the computational period. This performance is superficially inferior to that of the PPP model for the GDP-concept samples. However, the longer duration of the computational periods under the COL data set must be considered. The average computational periods for the COL samples are 6.5, 13.8, and 20.7 years, while the computational periods for the GDP-concept samples are 3, 5, and 20 years. Interestingly enough, for the COL data set, the superiority of the PPP over the naive model increases greatly when the computational period exceeds 10 years.

5.1.4 *Empirical Analysis of the Weak PPP and Naive Models*

An alternative PPP model, the "weak PPP model," involves a general linear relationship between P and D:

$$P = \alpha + \beta D + \in_3 \tag{5.6}$$

where α and β are parameters, with β positive. The corresponding naive model, designated as the "weak naive model," again ignores any information on relative price indexes in predicting the ratio of absolute PPP in the current period to absolute PPP in the base period. Therefore only the constant and the error term remain in Eq. (5.6). Thus, letting γ be a parameter, the weak naive model is as follows:

$$P = \gamma + \in_4 \tag{5.7}$$

The weak PPP model allows for a general linear relationship between P and D, with an error term. In contrast, the weak naive model specifies that P is equal to a constant (not necessarily unity) plus an error term. The implication of the naive model is that D can contribute nothing to the explanation of P. A significant correlation coefficient between P and D would indicate a linear association between the two variables and consequently forestall rejection of the weak PPP model.

The correlation coefficient (r) and coefficient of determination (r^2) between P and D for the six samples heretofore discussed are exhibited in Table 5.3. Because a negative correlation between P and D can be eliminated on theoretical grounds (according to the weak PPP model), a one-tail test of r is appropriate. Considering first the GDP-concept samples, the correlation coefficient for 1950–1955 is significantly different from zero at the 1% level. With only two degrees of freedom, the test of r for the 1967–1970 and 1950–1970 periods must be viewed with some skepticism. For neither of these samples is r significant at the 1% level; and while r is significant at the 5% level for the 1950–1970 sample, this result is largely due to an extreme observation, that for France.[3]

Table 5.3 Correlation of P with D

Sample	*Number of observations*	*Correlation coefficient*	*Coefficient of determination*
GDP concept			
1950–1955	8	0.93	0.86
1967–1970	4	0.92	0.84
1950–1970	4	0.98	0.96
Maximum-size	12	0.95	0.90
COL concept			
Less than 10 years	15	0.64	0.41
10 to 19 years	9	0.88	0.77
20 years or more	6	0.96	0.93
Maximum-size	22	0.90	0.81

Results are much better for the COL-concept samples. The correlation coefficient is uniformly significant at the 1% level. Indeed, even for a two-tail test, r continues to be significant at this level for the two longer-period samples and is significant at the 5% level for the less-than-ten-years sample.

With the objective of achieving more powerful tests of r than that provided by the individual samples, for each data set observations are pooled over these samples. While within the individual samples all observations are independent, this is not so between samples. Care must be taken to exclude observations that are dependent on other observations in the data set. The objective is to achieve a maximum-size sample of independent observations for each data set. For the GDP-concept data, this criterion involves excluding the 1950–1970 observations, resulting in a maximum-size sample of 12 observations. For the COL-concept data, 8 observations must be dropped from the longer computational periods, yielding a sample of 22 observations. Excluded observations are identified by superscript a in Tables 5.1 and 5.2.

The correlation coefficient and coefficient of determination for the maximum-size samples are presented in Table 5.3. For both samples, the correlation coefficient is significantly different from zero at the 1% level, even under a two-tail test—a result distinctly favorable for the weak PPP model.

The weak PPP and naive models have been tested against one another using correlation analysis; they may also be tested by means of regression analysis. The weak PPP model is formulated as Eq. (5.6), which can be viewed as a regression model the parameters of which, α and β, can be estimated by ordinary least-squares. The maximum-size samples from the GDP-concept and COL-concept data sets provide sample sizes (n) of 12 and 22 observations, respectively, to which Eq. (5.6) is fitted. The resulting regression lines are as follows:

$$\begin{array}{c} P = -.0666 + 1.0525D \\ (.1188) \quad (.1118) \end{array} \tag{5.8}$$

$$\begin{array}{c} P = -.4066 + 1.3948D \\ (.1699) \quad (.1487) \end{array} \tag{5.9}$$

The GDP-concept data yield Eq. (5.8) and the COL-concept data produce Eq. (5.9). Numbers in parentheses are standard errors of the estimated coefficients. The standard error of estimate (s) is 0.0322 in Eq. (5.8) and 0.1145 in Eq. (5.9), while the corrected coefficient of determination is 0.89 in (5.8) and 0.81 in (5.9).

The weak naive model is represented by Eq. (5.7), which may be viewed as the following regression model:

$$P = \gamma[1] + \in_4 \tag{5.10}$$

where [1] is a variable identically equal to unity. Equation (5.10) is a degenerate regression equation, the parameter of which, γ, may be estimated by ordinary least-squares. The resulting estimate of γ is the mean of the dependent variable ($\overline{P}$).

Letting $\hat{\alpha}$, $\hat{\beta}$, and "hatted" γ denote their estimates, parameter estimation of the weak PPP and naive models may be summarized as follows:

GDP-concept sample:

$$\hat{\alpha} = -.0666; \ \hat{\beta} = 1.0525; \ \hat{\gamma} = \overline{P} = 1.0488$$

COL-concept sample:

$$\hat{\alpha} = -.4066; \ \hat{\beta} = 1.3948; \ \hat{\gamma} = \overline{P} = 1.1709$$

Table 5.4 Summary of models

Model	*Value of* α	*Value of* β	F-*Statistic*	
			GDP concept	*COL concept*
Weak-PPP	$\hat{\alpha}$	$\hat{\beta}$	–	–
Weak-Naive	$\overline{P}$	0	44.24	44.03
Strong-PPP	0	1	0.81	4.86
Strong-Naive	1	0	58.07	68.51

Equation (5.6) may be used to exposit all four models that have been investigated empirically. Each of these models may be identified by the value it assigns to the parameters of Eq. (5.6), as shown in the second and third columns of Table 5.4. For the weak PPP and naive models, where the parameters lack preassigned numerical values, the estimates derived above are used.

The estimated regression Eqs. (5.8) and (5.9), which pertain to the weak PPP model, may be used for statistical testing of the remaining three models. As thus far only the weak PPP model has been subjected to testing with a level of significance (the correlation analysis), it is appropriate to use the estimated versions of this model for econometric testing of the other models.

The weak naive, strong PPP, and strong naive models each involve a joint hypothesis on α and β, the parameters of the weak PPP model. The respective hypotheses are indicated in the second and third columns of Table 5.4. For joint testing of α and β, the test statistic

$$F = \frac{n(\hat{\alpha} - \alpha)^2 + 2n\overline{D}(\hat{\alpha} - \alpha)\left(\hat{\beta} - \beta\right) + \sum D^2\left(\hat{\beta} - \beta\right)^2}{2s^2}$$

has the F-distribution with $(2,\ n - 2)$ degrees of freedom (Johnston 1972, pp. 28–29). The values of F computed for the $(\alpha,\ \beta)$ hypotheses implied by the weak naive, strong PPP, and strong naive models, respectively, are listed in the fourth and fifth columns of Table 5.4.

Results are most striking for the GDP-concept sample. With (2, 10) degrees of freedom, the critical value for the F-distribution at the 0.05% level of significance is 17.9. Both naive models are rejected at this extremely low level of significance. In contrast, the F value for the strong

PPP model is such that this model cannot be rejected even at the 40% level of significance.[4] Thus, using the *F*-test as the criterion, the strong PPP model strongly out-performs both naive models.

For the COL-concept sample, the *F*-distribution has (2, 20) degrees of freedom; its critical value at the 0.05% level of significance is 11.4. Again the naive models are rejected at this very low level of significance. Indeed, the *F*-statistics for these models are even further in the tail of the distribution than under the GDP-concept sample. A statement similar in kind applies to the *F*-statistic of the strong PPP model. With *F*-distribution critical values of 3.49 and 5.85 at the 5% and 1% levels, respectively, the strong PPP model itself is rejected at the 5% level of significance, though it cannot be rejected at the 1% level. So the strong PPP model survives the *F*-test more easily under the GDP-concept sample than under the COL-concept sample.

In summary, empirical investigation supports the relative PPP hypothesis in both strong and weak versions. The findings are especially favorable to the relative PPP theory in light of the facts that (i) no secondary variables were used to increase the explanatory power of the PPP model, (ii) complicated functional forms and lagged relationships were not adopted in an effort to increase explanatory power, and (iii) several countries outside the Western industrial mode were included in the samples.

Appendix: The Data

Absolute PPP

The absolute PPP measure is computed as the geometric mean of the PPPs calculated alternatively using the weighting pattern of the domestic country and that of the standard country. If the weights of only one of the countries are used for the PPP computation, then the calculated PPP will be biased in the direction of an overvalued PPP for that country (Houthakker 1962, p. 297; Officer 1976, pp. 15–16). Therefore no use is made of PPP measures for which only one of the weighting patterns is available.

Two data sources are used to obtain absolute-PPP measures under the GDP concept: Gilbert and associates (1958), who provide PPP data for the years 1950 and 1955, and Kravis et al. (1975), who offer such data for 1967 and 1970. The former authors use a GNP rather than GDP price-level concept. Because the PPP theory concerns prices and production within the boundaries of respective countries, GDP is the preferred concept, as it covers domestic rather than national production. In practice, PPP computations on a GNP basis differ minimally from those on a GDP basis.

The PPP data published by the German Statistical Office (Statistisches Bundesamt) are the source of the COL-concept measures.

GDP Deflator

The data source is OECD national-accounts publications for all countries except Hungary and Kenya, and United Nations (1973) for the latter two countries. For Hungary, GDP data are not available and a less-inclusive concept of domestic production, "net material product," must be used.

Consumer Price Index

For all countries except the Soviet Union, the data source is International Monetary Fund (1977). For the Soviet Union, use is made of the "state retail price index," published in International Labour Office (1962).

Notes

1. The term "standard" is used in preference to "foreign" currency or country, because this country may serve as the standard of comparison for a group of "domestic" countries.
2. Details on data are provided in the appendix.
3. See Table 5.1. With the value of P extremely close to that of D, coupled with by far the highest values of both these variables in the 1950–1970 sample, the observation for France is dominant in the correlation.
4. Critical values for the F-distribution at the 40% level of significance are not published in an accessible source. However, the critical value at this level for the F-distribution with $(2, \infty)$ degrees of freedom is readily obtained as 0.916; for this degrees-of-freedom configuration reduces the F-distribution to a chi-square distribution with 2 degrees of freedom. The corresponding critical value for the F-distribution with (2, 10) degrees of freedom is necessarily greater than 0.916.

References

Gilbert, Milton, and Associates, *Comparative National Products and Price Levels* (Paris: Organization for European Economic Cooperation, 1958).

Houthakker, Hendrik S., “Exchange Rate Adjustment,” in *Factors Affecting the United States Balance of Payments*, Compilation of Studies Prepared for the Subcommittee on International Exchange and Payments, Joint Economic Committee, 87th Congress, 2nd Session (Washington, DC: U.S. Government Printing Office, 1962), 287–304.

International Labour Office, *Yearbook of Labour Statistics 1962* (Geneva, 1962).

International Monetary Fund, *International Financial Statistics*, 1977 Supplement.

Johnston, J., *Econometric Methods* (New York: McGraw-Hill Book Company, 1972).

Kravis, I.B., Kenessey, Z., Heston, A., and R. Summers, *A System of International Comparisons of Gross Product and Purchasing Power* (Baltimore: Johns Hopkins University Press, 1975).

Officer, Lawrence H., “Purchasing Power Parity and Factor Price Equalization,” *Kyklos* 27 (Fasc. 4, 1974), 868–878.

______, “The Purchasing-Power-Parity Theory of Exchange Rates: A Review Article,” *International Monetary Fund Staff Papers* 23 (March 1976), 1–60.

Organization for Economic Cooperation and Development (OECD), *National Accounts of OECD Countries*, various issues.

Statistisches Bundesamt, *Preise, Löhne, Wirtschaftsrechungen, Reiche 10: Internationaler Vergleich der Preise für die Lebenshaltung*, various issues.

United Nations, *Yearbook of National Accounts Statistics*, volume I (1973).

CHAPTER 6

Law of One Price

6.1 The Law of One Price Cannot Be Rejected: Two Tests Based on the Tradable/Nontradable Price Ratio

Republished with permission of Elsevier, from *Journal of Macroeconomics*, Vol. 8, No. 2 (Spring 1986), pp. 159–182; permission conveyed through Copyright Clearance Center, Inc.

6.1.1 Introduction

The law of one price for tradable commodities is an essential ingredient in the body of knowledge known as international economics. Without the imposition of this law, there would not even be the traditional "pure theory" of international trade. Without this law, much of the "monetary theory," too, would have to be reconstructed.

Yet, the empirical evidence does not support the law of one price. On the contrary, with the exception of tests involving narrowly defined, extremely homogeneous commodities, the law has been universally rejected in econometric and other testing.

The premise of this paper is that the reason for the failure of the law of one price in empirical testing is the disaggregative approach almost uniformly followed. An aggregative technique is superior on several

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_6

grounds: theoretically, because it accounts for cross-commodity substitution in production and consumption; statistically, because it avoids the difficult task of matching individual products across countries; and econometrically, because it commits no specification errors.

Therefore, rather than striving for commodity homogeneity among countries, I use a broad sectoral classification of tradables versus nontradables. An equation is developed involving purchasing power parity (PPP), the exchange rate, and the nontradable/tradable price ratio in each country. This relationship is used to perform two novel tests of the law of one price: first, equality of the prices of tradables and nontradables for a given country; second, equality of the price of tradables across countries. If the first test is successful, then the law of one price for tradables can be extended to a law of one price for all commodities, both tradables and nontradables.

In Sect. 6.1.2 the law of one price is formally expressed and the reasons for deviations from the law discussed. Section 6.1.3 surveys the empirical literature on the law of one price, and Sect. 6.1.4 generates the model to be used in the present, aggregate, testing of the law. Using data developed in Sect. 6.1.5, the equality of the prices of tradables and nontradables is tested in Sect. 6.1.6 and equality of the price of tradables across countries in Sect. 6.1.7. Tests of the validity of the aggregative approach are described in Sect. 6.1.8, and concluding comments of the study are made in Sect. 6.1.9, followed by an appendix on the data.

6.1.2 Law of One Price

The law of one price for tradables states that there is a unique price of a tradable commodity irrespective of the country of output, where the respective home-currency prices of the commodity are expressed in a common currency via market exchange rates. If one abstracts from the inevitable index-number problems of aggregation, this law, if applicable at the disaggregative level, would farther hold for aggregates of tradable commodities, in particular, the totality of tradables. For the law of one price to be extended from tradables to all commodities, one requires the further relationship that, for each country, the price level of tradables is equal to that of nontradables.

For the law of one price of tradables to be valid, a sufficient condition is that the markets involved be purely and perfectly competitive (in the Chamberlinian sense). This would assure the existence of perfect arbitrage. Further, if the commodities in a market comparison are not identical, elasticities of substitution in production and/or consumption must nevertheless be high. To extend the validity of the law to nontradables, tradables and nontradables should again have substantial substitution possibilities with respect to each other in production and/or consumption.

Looked at from the opposite standpoint, what are the elements that give rise to deviations from the law of one price? These forces are threefold. First, the purity of competition may be lacking. The existence of monopoly and oligopoly can cause divergences from the law for two reasons: (i) the monopolist may practice price discrimination in the domestic and foreign markets (see Ripley 1974; Goldstein and Officer 1979; Crouhy-Veyrac et al. 1982); (ii) oligopolists, in fulfillment of a desire for price stability, may absorb the impact of a changing exchange rate in their profits, so that the price of tradables does not move with the exchange rate to maintain the law of one price (see Dunn 1970, 1973). Second, the phenomenon of product differentiation can reduce the substitutability of manufactured goods of different countries, even for products within the same commodity category (see Kravis and Lipsey 1971, 1978; Norman 1975; Isard 1977b). Third, at an aggregate level, the price of tradables may have differing weighting patterns (that is, differing commodity compositions) in the countries involved (see Isard 1977a; Kravis and Lipsey; 1978; Goldstein and Officer 1979; Crouhy-Veyrac et al. 1982).

6.1.3 Critique of Empirical Literature

Like all economic theories, the law of one price can hold only to an approximate degree in the real world. Still, even while allowing for random errors as well as for the systematic factors making for deviations from the law, the law of one price should be expected to hold empirically to a considerable extent; for its foundations—competitive conditions and high elasticities—are the basic requisites of a well-functioning economic system. In this light, it would be surprising—and demoralizing for the domestic and international economy—if the law of one price tended to be rejected by the evidence. And yet, incredibly, of the sixteen empirical studies on the issue of which I am aware, thirteen have negative implications for the law of one price.[1] Of the remainder, two (Genberg 1975; Rosenberg 1977) can be construed as supporting the law of one price; while the third (Crouhy-Veyrac et al. 1982) validates the law only for what the authors call "two pilot cases" as distinct from "the main evidence." Indeed, only for primary products priced on international commodity exchanges has the law of one price received consistent validation (by Genberg and Crouhy-Veyrac and others). In the remaining study (Rosenberg) in which the law of one price receives support, the commodities are precisely defined steel products made homogeneous now by manufacturing design rather than by nature.

When products have any differentiation at all, the law of one price ceases to receive support. It is my contention that the reason is the decision to compare commodities at disaggregative levels, this decision made by all the authors involved. Only one existing test of the law of one price is at the level of aggregate tradables. Kravis and Lipsey (1978, p. 222) conclude that differences in tradables prices are "not trivial even among the industrial countries." However, the range of the price level for tradables is less than half that for nontradables, suggesting that this test, on balance, is not that unfavorable regarding the law of one price.

It is the position of this paper not only that the appropriate level at which to test the law of one price is that of aggregate tradables (and nontradables) but also that a formal model should be used to derive an estimable equation, permitting comparison of results "with versus without" imposition of the law of one price (the latter not performed by Kravis and Lipsey). In contrast, the conventional treatment is commodity disaggregation combined with comparisons via ad hoc observation or simple correlation or regression analysis. The conventional approach has several weaknesses that bias results against the law of one price:

(i) The law of one price in practice might emanate from more complex substitutions in production and consumption than those of a bilateral nature, that are inherent in disaggregative testing. The complex substitutions, not captured in disaggregative testing, are incorporated in an aggregative approach such as that of the present paper. Further, because the conventional, disaggregative approach involves testing for purely bilateral, one-on-one substitution, the data must be made commensurate across countries: each product examined must be made homogeneous or comparable over all sources of supply—an extremely difficult task given both the nature of many manufactured goods and the specific peculiarities of each country's official statistics. This problem does not at all exist in the aggregative approach.

(ii) The existence of transport, insurance, information, and other transactions costs, together with trade restrictions, and of changes in these elements, implies that conventional disaggregative testing (as well as the aggregate-level testing carried out by Kravis and Lipsey 1978) is too severe on the law of one price.[2] Testing the null hypothesis of unity for the common-currency domestic/foreign price ratio of a commodity biases the finding in favor of rejecting the law. Rather, each of the end points, or "commodity points," delimited by the transactions costs and trade restrictions or their changes and within which the law of one price is valid, must be the subject of the null hypothesis. This procedure has not been followed in any testing to date. The problem itself does not even arise in the aggregate-level testing of the present study, as the "null hypothesis" is tested not directly but only indirectly.

(iii) Those authors that regress the domestic country's price index on the foreign price index and the exchange rate as separate independent variables are probably committing a specification error. The reasons why there might be a different response to a change in the one rather than the other explanatory variable are quite subsidiary to the main issue of the law of one price (see Crouhy-Veyrac et al. 1982, pp. 331–332), Empirically, Crouhy-Veyrac and others (1982) find that decomposing the explanatory variable worsens their regressions considerably. These arguments and findings suggest that the negative results of Curtis (1971), Bordo and Choudhri (1976), Kravis and Lipsey (1977), and Richardson (1978) are all suspect, as these authors specify a decomposition of the explanatory variable.

6.1.4 A Model of PPP and the Tradable/Nontradable Price Ratio

6.1.4.1 Derivation of PPP/Exchange-Rate Relationship

In this section a model is developed to test the law of one price in a way quite different from the traditional approach. The totality of production, or gross domestic product (GDP), is divided into two categories, tradables and nontradables, for both a domestic country (i) and a base country (b). The purchasing power parity, or relative price levels, of the countries is defined in terms of the prices of tradables and nontradables:

$$\mathrm{PPP}^i \equiv \frac{\mathrm{WT}^i \cdot \mathrm{PT}^i + \mathrm{WN}^i \cdot \mathrm{PN}^i}{\mathrm{WT}^b \cdot \mathrm{PT}^b + \mathrm{WN}^b \cdot \mathrm{PN}^b} \tag{6.1}$$

where

$$\mathrm{WT}^j + \mathrm{WN}^j \equiv 1 \quad j = i, b \tag{6.2}$$

with the following notation:

PPP^i = purchasing power parity for country i, no. of units of i's currency per unit of base currency
PT^j = price level of tradables in country j, with a weight of WT^j in the overall price level; $j = i, b$
PN^j = price level of nontradables in country j, with a weight of WN^j in the overall price level; $j = i, b$
R^i = exchange rate for country i's currency, no. of units of i's currency per unit of base currency.

PPP theory suggests that a country's own production pattern is the optimal weighting scheme for its price level, and the country-specific weights in Eq. (6.1) reflect this fact.[3] Each country's price level is a weighted average of commodity prices in the country, with own-country production (expenditure) weights.

At the adopted level of aggregation the law of one price for tradables is:

$$\mathrm{R}^i = \mathrm{PT}^i / \mathrm{PT}^b \tag{6.3}$$

Combining Eqs. (6.1) and (6.3), one obtains:

$$\mathrm{PPP}^i \Big/ R^i \equiv \frac{\mathrm{WT}^i + \mathrm{WN}^i \cdot \left(\mathrm{PN}^i \Big/ \mathrm{PT}^i\right)}{\mathrm{WT}^b + \mathrm{WN}^b \cdot \left(\mathrm{PN}^b \Big/ \mathrm{PT}^b\right)} \tag{6.4}$$

subject, of course, to Eq. (6.2). Equation (6.4) then involves the purchasing power parity, the exchange rate, the nontradable/tradable price ratio in the two countries ($\mathrm{PN}^i/\mathrm{PT}^i$ and $\mathrm{PN}^b/\mathrm{PT}^b$), and the weight of tradables in the countries' respective price levels.

6.1.4.2 Generation of Estimable Equation

Thus far the analysis has dealt with a single time period. The problem with Eq. (6.4), therefore, is that the two nontradable/tradable price variables are ratios of price levels rather than of period-to-period indexes and therefore are nonobservable; nor have such data been constructed except in normalized form, rendering them useless for Eq. (6.4) (see, for example, Kravis et al. 1982, pp. 193–196). For an estimable equation, one can have this nonobservable variable just for the base country; for there is only one such country, while i ranges over N domestic countries, with N the sample size.

The unknown $\mathrm{PN}^i/\mathrm{PT}^i$ is eliminated by considering two time periods, a "current period," t, and "base period," o, respectively subscripting variables as such, rearranging Eq. (6.4), and taking the ratio of the equations in the two periods. Then, after considerable algebraic manipulation, one obtains:

$$\left(\mathrm{PPP}^i/R^t\right)_t = \frac{A^i + B^i \cdot \beta}{C^i + D^i \cdot \beta}, \tag{6.5}$$

where

$$A^i \equiv \left(\mathrm{WT}^i/\mathrm{WN}^i\right)_t + I^i\left[\left(\mathrm{PPP}^i/R^i\right)_o\left(\mathrm{WT}^b/\mathrm{WN}^i\right)_o - \left(\mathrm{WT}^i/\mathrm{WN}^i\right)_o\right],$$

$$B^i \equiv \left(I^i/I^b\right)\left(\mathrm{PPP}^i/R^i\right)_o\left(\mathrm{WN}^b/\mathrm{WN}^i\right)_o,$$

$$C^i \equiv \left(\mathrm{WT}^b/\mathrm{WN}^i\right)_t,$$

$$D^i \equiv \left(\mathrm{WN}^b/\mathrm{WN}^i\right)_t,$$

$$I^j \equiv (\mathrm{PN}/\mathrm{PT})_t^j/(\mathrm{PN}/\mathrm{PT})_0^j \quad j = i, b,$$

$$\beta \equiv (\mathrm{PN}/\mathrm{PT})_t^b.$$

Noting that the I^j variables are ratios of period-to-period nontradable/tradable price ratios (or, equivalently, ratios of a nontradable price index to a tradable price index), these variables are observable or, more precisely, calculable. In fact, data can be obtained for all variables in Eq. (6.5) with the exception of $\left(\mathrm{PN}/\mathrm{PT}\right)_t^b$, or β, the nontradable/tradable price ratio in the base country. The first test of the law of one price is now apparent. One specifies an error structure for Eq. (6.5), a base country, b, and a current period, t, and assembles data to construct the variables A^i, B^i, C^i, D^i, I^i, l^b for a sample of N domestic countries, that is, $i = 1, \ldots, N$.[4] The parameter β is then estimated econometrically from this observation matrix. Assuming that the law of one price for tradables [Eq. (6.3)] holds, this parameter is the ratio of the price level of nontradables to that of tradables. If its estimate is not significantly different from unity, while significantly different from zero, then tradables and nontradables for a given country (or at least for the base country) are good substitutes; it could not be rejected that a law of one price for tradables widens to a law of one price for commodities generally.

The second test of the law of one price is of the equality of the price of tradables across countries, that is, whether Eq. (6.3) holds empirically. The technique is to drop Eq. (6.3) from the model, thus permitting a comparison of results including versus excluding the law of one price for tradables.

6.1.4.3 Limitations of PPP Concept

The testing procedure thus described, while devoid of the limitations of the disaggregative approach, is not without its costs. In order to obtain an estimable Eq. (6.5), PPP had to be defined in an unconventional way and in two respects. The weights for the countries' price levels are not only country-specific but also expenditure rather than quantity based. While the own-country weighting is justified in terms of PPP theory, the two weighting properties together imply that purchasing power parity, the countries' relative price levels [Eq. (6.1)], is not a true price index in the sense that it can be re-expressed as a meaningful function of price relatives. It is possible, therefore, that PPP as defined is sensitive to a change in the unit of measurement of a commodity. Fortunately, it is arguable that the problem is not serious, in part because it is legitimate to impose the base country's unit of measurement in the base period on all countries and both periods. Then a proportionate change in measurement units across all commodities would not affect the PPPs. Also, the PPPs are now single-valued, though only with respect to the customary measurement units of the base country. The extent of residual ambiguities in the PPP concept would be an empirical question, but one that is irrelevant for this study. The reason is that PPPs defined as in Eq. (6.1) are unobservable and resort must be had to conventional PPPs that are normal price indexes.

Similar limitations apply to the ratio of the price level of nontradables to that of tradables within a country, and hence to testing for equality of the prices of tradables and nontradables. In this case, the issue is resolved through the ratio appearing only as an estimable parameter in the model.

6.1.5 The Sample

6.1.5.1 Selection of PPP Measure

With PPP defined as in Eq. (6.1) unavailable, Irving Fisher's ideal index number is selected as the PPP measure. As the geometric mean of the Laspeyres (base-country-weighted) and Paasche (domestic-country-weighted) indexes, the Fisher index has the property of "equi-characteristicity," that is, equal consideration is given to the weighting pattern of each country. In contrast, the Laspeyres and Paasche indexes each have the well-known bias of a relatively lower price level for the country whose weights are used. Looked at another way, Eq. (6.1) defines the price level for each country in terms of its own weighting pattern. The Laspeyres and Paasche indexes each satisfy this criterion for one country and contradict it for the other; the logical compromise is Fisher's ideal index.

6.1.5.2 Selection of PN/PT Measure

Consider Eq. (6.1), which defines the overall price level for each country as a weighted average of its tradable and nontradable components, PT and PN, respectively. What better measure of the overall price level could there be than the GDP price level, which aggregates the prices of all domestic production, that is, the total of tradable and nontradable output, with weights proportional to domestic output of tradables and nontradables, respectively? Switching to index numbers, the GDP deflator, PGDP, is the same weighted average of price-index equivalents of PT and PN. Since PGDP is constructed as the ratio of current-priced to constant-priced aggregate output (GDP), the price deflators for tradable and nontradable output—denoted as $\hat{\mathrm{P}}\mathrm{T}$ and $\hat{\mathrm{P}}\mathrm{N}$, respectively—are obtained along the same lines, with the output of tradables (nontradables) defined as that part of GDP originating in the tradable (nontradable) sector of the economy. It remains only to allocate industries to the tradable and nontradable sectors. The tradable sector is taken to consist of (1) agriculture, hunting, forestry, and fishing, (2) mining and quarrying, and (3) manufacturing, while the nontradable sector is composed of all other industries in which GDP originates.[5] Then the nontradable/tradable price-index ratio, I^J, is $\hat{P}N \big/ \hat{P}T$.[6] This variable provides an obvious and logically compatible nontradable/tradable weighting pattern (WN, WT), namely, the proportion of constant-priced output originating in the (nontradable, tradable) sector.

6.1.5.3 Selection of Sample

The sample size is delimited by the availability of two sets of data: PPP measures (to construct the PPP/*R* variables) and output by industry of origin (to construct the $\hat{P}N \big/ \hat{P}T$ variables). Available data on economy-wide PPP indexes are most extensive for the United States as the country of comparison; so it is the logical choice for base country. Restricting the search for PPP data to those at an economy-wide level (GDP, GNP, and NNP, in that order of preference) and that are Fisher indexes with the United States as the common base country, 1975 is the year ("current year") for which by far the largest number of data points, at 34, is obtainable. Of these countries, eighteen have at least one other year (a "base year") of PPP data satisfying the above criteria, while four fulfill the criteria in all respects except one: the foreign country against which comparison is made is not the United States. These four remain in the sample by obtaining the base-year PPP via a linking process (see data appendix). The other twelve countries are dropped from the sample, and two further countries are eliminated because of a lack of availability of tradable/nontradable data to construct the $\hat{P}N \big/ \hat{P}T$ variable, resulting in a final sample size of twenty.

Several countries in the sample have more than one base year for which both the PPP and tradable/nontradable data are available. A unique base year is obtained by selecting the one furthest in the past, in order to provide a maximum time span over which to test the law of one price. The resulting 20-country sample is summarized in Part A of Table 6.1, with the country, base year, and value of the PPP/*R* variable in base and current periods shown in the first four columns.

Table 6.1 PPP/R data and residuals from its estimates

		PPP/R		*Residuals*[b]		
Country	*Base year*	*Current year*[a]	*Base year*	*Equation*	*PGDP method*	*Naive model*
A. Sample size of 20						
U.K	1950	0.8627	0.7114	0.0199	0.1498	−0.1513
Belgium	1955	1.1469	0.8182	−0.1368	0.0962	−0.3287
Denmark	1950	1.2909	0.7117	−0.3477	0.2357	−0.5793
France	1950	1.1028	0.7545	−0.1041	0.0022	−0.3483
Germany	1950	1.1706	0.7318	−0.2609	0.3517	−0.4388
Italy	1950	0.8657	0.6980	0.0830	0.2063	−0.1677
Netherlands	1970	1.1811	0.7770	−0.1437	0.0828	−0.4041
Hungary	1967	0.6041	0.7453	0.1009	0.0150	0.1412
Poland	1965	0.7716	0.7333	0.0090	−0.1444	−0.0383
Japan	1967	0.9364	0.6544	−0.0436	0.0219	−0.2819
Brazil	1968	0.6642	0.5124	−0.0195	−0.0853	−0.1518
Colombia	1970	0.4054	0.4392	0.2040	0.0442	0.0337
Mexico	1968	0.5736	0.4884	0.0545	0.0519	−0.0852
Uruguay	1968	0.5485	0.3957	0.0031	−0.0407	−0.1528
Iran	1970	0.6660	0.4779	−0.0130	0.2968	−0.1882
Kenya	1967	0.5519	0.5502	0.1283	0.0166	−0.0017
India	1967	0.3364	0.3609	0.0875	0.0146	0.0245
Korea	1970	0.4537	0.5635	0.3052	0.1686	0.1098
Philippines	1970	0.4383	0.4013	0.1456	0.0378	−0.0370
Thailand	1963	0.4510	0.2489	−0.0717	−0.1798	−0.2021
B. Sample size of 2						
Norway	1950	0.7858	0.6825	−0.0302	0.0012	−0.1033
Canada	1950	0.9072	0.8476	0.0324	−0.0139	−0.0597

[a]Year of predicted PPP/*R*. 1975 for all countries in 20-country sample; 1955 for Norway and 1965 for Canada
[b]Estimated *minus* actual PPP/*R*

6.1.6 Test of Equality of Prices of Tradables and Nontradables

6.1.6.1 Estimation Technique

All variables in Eq. (6.5) are observable, with the exception of $(\mathrm{PN}/\mathrm{PT})_t^b$ the price-level ratio in the base country in the current period, that is, the parameter β. Assuming an additive error term that is independently and identically distributed for all observations, nonlinear least-squares is an appropriate method, providing a consistent and asymptotically normally distributed estimate of β, and this is the technique adopted.[7]

Table 6.2 Estimates of regression equation

Equation number	$\hat{\alpha}$	$\hat{\beta}$	*Correlation coefficient*[a]	*Log-likelihood ratio*[b]
1	0.17 (4.59)	0.84 (3.96)	0.89	43.69
2	–	1.24 (2.47)	0.86	35.88

[a]Correlation of actual and fitted dependent variable
[b]Restriction $\beta = 0$

6.1.6.2 *Estimation of Equation (6.5)*

Because the countries composing the sample do not have a common base period, a better fit of Eq. (6.5) could be obtained by including a constant term, α, and the resulting regression is presented as Equation number 1 in Table 6.2, where $\hat{\alpha}$ and $\hat{\beta}$ are the nonlinear least-squares estimates of α and β, respectively, with their t-values in parentheses. Now, $\hat{\beta}$ is the estimate of $(\mathrm{PN}/\mathrm{PT})_t^b$, the nontradable/tradable price-level ratio in the United States in 1975. According to two statistics—the t-test (with values 3.96 and -0.74) and the log-likelihood-ratio test (with values 43.69 and 0.48)—β is both significantly different from zero at extremely low levels of significance (less than one tenth of one percent) and not significantly different from unity at extremely high levels of significance (above 40%). Furthermore, the point estimate of β, at 0.84, is itself not far away from unity. Therefore one cannot reject the hypothesis that the price levels of tradables and nontradables were equal for the United States in 1975.

Division of $\hat{\beta} = 0.84$, the estimated U.S. nontradable/tradable price-level ratio in 1975, by the corresponding U.S. price-index ratio, I^{b}, for "base periods" at five-year intervals between 1950 and 1980 yields a time series of the U.S. nontradable/tradable price-level ratio, PN/PT, as shown in the second column of Table 6.3. While historically PN/PT has been rising, by 1970–1980 it stabilized not far from unity.

6.1.6.3 *Performance of Equation Outside of Sample*

There are two countries, Norway and Canada, for which appropriate PPP data, while not available for 1975, exist for two or more other years.[8] The requisite tradable/nontradable data are also available, resulting in a two-country sample outside the original sample and exhibited in Part B of Table 6.1 (first four columns). It would be inappropriate to use Equation number 1 to predict "current-period" PPP/R for Norway and

Table 6.3 Time series of U.S. nontradable/tradable price-level ratio

Year	*Based on*	
	Equation 6.1	*Equation 6.2*
1950	0.63	0.93
1955	0.67	0.98
1960	0.69	1.02
1965	0.77	1.13
1970	0.85	1.25
1975	0.84	1.24
1980	0.84	1.23

Canada in 1955 and 1965, respectively (their current periods), because that regression has a constant term predicated on the current year 1975 and a specific conglomeration of base years. The better procedure is to drop the constant and re-estimate the equation, with the result exhibited as Equation number 2 in Table 6.2. Applying the U.S. nontradable/tradable price-index ratio, l^b, to $\hat{\beta} = 1.24$, as was done above for $\hat{\beta} = 0.84$, one obtains another time series of the U.S. absolute price-level ratio, PN/PT, as shown in the third column of Table 6.3. Plugging in the appropriate values of A^i, B^i, C^i, D^i, l^i, I^b for $i =$ (Norway, Canada), $t =$ (1955, 1965), and $o =$ (1950, 1950), where $\hat{\beta} =$ (0.98, 1.13), the result is the predicted PPP/R in period t via the equation. The forecast error (difference between the estimated and actual PPP/R) is shown in column 5 of Table 6.1 (last two rows, for Norway and Canada). The absolute error amounts to only 3.84% of the true PPP/R for Norway and 3.57% for Canada.

6.1.7 Test of Equality of Prices of Tradables Across Countries

6.1.7.1 Motivation of Test Procedure

The law of one price for tradables is tested by considering Eq. (6.5) as forecasting the dependent variable $(\mathrm{PPP}^i/R^i)_t$, the PPP/exchange-rate ratio for the domestic country in 1975. As Eq. (6.5) embodies the law of one price [Eq. (6.3)], an alternative predictor of $(\mathrm{PPP}^i/R^i)_t$ is obtained by dropping Eq. (6.3) from the model. If the estimates from the two predictors—the first of which does, the second of which does not incorporate the law of one price—are sufficiently close, then the law of one price cannot be rejected. As an indicator of closeness, a third predictor of $(\mathrm{PPP}^i/R^i)_t$ is developed, based on a naive model. For the law of one price to be supported, Eq. (6.5)'s prediction of PPP/R must be close to the estimate of the model excluding Eq. (6.3) but far from the prediction of the naive model.

*6.1.7.2 Alternative Methods of Predicting PPP/*R

Equation (6.5) is interpreted as predicting the PPP/R ratio in period t given the ratio in period o and subject to the law of one price, Eq. (6.3), holding in both periods. To set up a test of the law of one price, Eq. (6.3) is dropped but the rest of the model retained. Consider the GDP deflator defined in terms of the prices of tradables and nontradables:

$$\mathrm{PGDP}_t^j \equiv \frac{\mathrm{WT}_t^j \cdot \mathrm{PT}_t^j + \mathrm{WN}_t^j \cdot \mathrm{PN}_t^j}{\mathrm{WT}_t^j \cdot \mathrm{PT}_o^j + \mathrm{WN}_t^j \cdot \mathrm{PN}_o^j} \quad j = i, b \tag{6.6}$$

where

$\mathrm{PGDP}^j{}_t$ = GDP deflator for country j in period t relative to period o $j = i, b$

Then $\mathrm{PPP}^i{}_t$ may be approximated by

$$\left(\mathrm{PGDP}_t^i/\mathrm{PGDP}_t^b\right) \cdot \mathrm{PPP}_o^i \tag{6.7}$$

and $(\mathrm{PPP}^i/R^i)_t$ estimated as

$$\left[\left(\mathrm{PGDP}_t^i/\mathrm{PGDP}_t^b\right) \cdot \mathrm{PPP}_o^i\right]/R_t^i \tag{6.8}$$

Note that, as the GDP deflator is a current-weighted price index, expressions (6.7) and (6.8) are only approximations to $\mathrm{PPP}^i{}_t$ and $(\mathrm{PPP}^i/R^i)_t$, respectively. Were the denominator of the right-hand side of Eq. (6.6) to involve $(\mathrm{WT}^j{}_o, \mathrm{WN}^j{}_o)$ in place of $(\mathrm{WT}^j{}_t, \mathrm{WN}^j{}_t)$, then (6.7) and (6.8) would be identically equal to $\mathrm{PPP}^i{}_t$ and $(\mathrm{PPP}^i/R^i)_t$.

This approach to estimation of the PPP/R ratio in period t may be called the "PGDP" method in opposition to the "equation" method, based on Eq. (6.5). The PGDP method is used to obtain predictions of PPP/R in the current period for each of the twenty countries in the original sample and the two countries in the second sample. Defining the residual as the difference between the estimated and actual PPP/R, this error is listed for the equation method (that is, the estimates via Equation numbers 1 and 2 (in Table 6.2) for the twenty-country and two-country samples, respectively) and the PGDP technique in columns 5 and 6, respectively, of Table 6.1. Which method could be expected a priori to lead to the better forecast? To answer the question, consider the various sources of error.

(i) According to Eq. (6.1), the price levels composing PPP are to be defined with country-specific weights. Instead, a compromise measure, the Fisher index, is used. This substitution applies to both techniques.
(ii) Several data problems also affect both approaches. Since the GDP deflator is obtained in practice as the ratio of current-priced to constant-priced GDP, any change in the base period of the constant-priced series (even though corrected via linking on the basis of an overlap) will disturb the PGDP series and therefore the PGDP-method prediction. So will a switch to a new system of national accounts not carried back to base period o. The same data problems apply to the equation method, however, because crucial to the method is the computation of tradables and nontradables "deflators" using national-accounts data. Further, the PPP and PGDP series may not be comparable conceptually; but again the same caveat applies to the PPP and nontradable/tradable price-index series.
(iii) The PGDP method assumes that the weights (WT, WN) for the current period apply also to the base period. The equation method allows for differences in the (WT, WN) weights in the two periods. This is an advantage of the equation approach only if the tradable/nontradable division of output that it adopts is sufficiently consonant with reality.
(iv) As a direct result of imposing the law of one price, the equation technique explicitly incorporates the nontradable/tradable price-index ratio in the domestic and base countries. Nothing is gained in prediction by this complication. There is an unnecessary complexity in the equation's forecast compared to the PGDP predictor. On the average, one would expect such complexity to bring about a greater magnitude of forecast error.

Table 6.4 Average errors from estimates of PPP/R

Sample size	*Percent of mean true PPP/*R					
	Average of absolute errors			*Average of squared errors*		
	Equation method	*PGDP method*	*Naive model*	*Equation method*	*PGDP method*	*Naive model*
20	15.19	14.93	25.74	2.98	2.97	8.11
2	3.70	0.89	9.63	0.12	0.01	0.84

Interestingly, the equation method has the smaller residual (in absolute value) for as many as 8 of the 22 forecasts, as shown in columns 5 and 6 of Table 6.1. The comparison is suggestive that the law of one price cannot be rejected. The conclusion is reinforced by computation of the average absolute errors and average squared errors (each expressed as a percentage of the mean true PPP/R) in Table 6.4. For the twenty-country sample, the error level is relatively high and the equation technique has average error extremely close to the PGDP method. For the two-country sample, the gap between the errors is wider, but the level of the errors is much smaller. So the errors resulting from the two methods appear to be quite close. Still, "how close is close?"

6.1.7.3 Comparison with Naive Model

Consider yet a third method of estimating $(\text{PPP}^i/R^i)_t$—a naive model that ignores all price changes between the base and current period and predicts PPP_t by PPP_o, thus estimating $(\text{PPP}^i/R^i)_t$ as $\text{PPP}^i{}_o/R^i{}_t$.

The residuals from the naive model are listed in the final column of Table 6.1 and the resulting average errors presented in Table 6.4 along with those of the other two techniques. On the average, for each sample, the residuals rank as follows (and as expected) in absolute value (smallest first): PGDP method, equation method, naive model. For each average-error measure (absolute and squared) and each sample, define an "error-difference ratio" as follows:

(Equation-Method Error *minus* PGDP-Method Error)/(Naive-Model Error *minus* Equation-Method Error).

Table 6.5 Error-difference ratios

Sample size	*Based on*			
	Absolute errors		*Squared errors*	
	Ratio	*Inverse*	*Ratio*	*Inverse*
20	0.02	40.58	0.002	513.00
2	0.47	2.11	0.15	6.55

This ratio provides a heuristic test of the law of one price. If its value is unity, the equation average residual is equidistant between those of the other two techniques, providing neither positive nor negative evidence for the law of one price. If the ratio exceeds unity, this has a negative implication for the law of one price. Below unity, the law of one price is supported. The lower the value of the ratio, the greater the evidence for the law of one price. Taking the inverse of the ratio, the significance of deviations from unity is reversed. Now the higher the ratio (providing it is above unity), the more support there is for the law.

Table 6.5 presents the error-difference ratios and their inverses. While the test itself has no statistical significance, the ratios are sufficiently far below unity (or, the inverse ratios sufficiently far above unity)—especially for the twenty-country sample—that the following conclusion can reasonably be drawn: the law of one price cannot be rejected on the basis of the evidence presented in this study.

6.1.8 Validity of Tradable/Nontradable Distinction

If the allocation of goods into the tradable and nontradable sectors is largely arbitrary, then the evidence in favor of the law of one price becomes suspect. Suppose that the tradable/nontradable distinction involves little more than a random allocation of commodities into the two sectors. The implication for the finding that prices of tradables and nontradables are equal is that, with each sector receiving a random allocation of goods, of course their price levels would be approximately the same. However, the result that prices of tradables are equal across countries is supportive of the law of one price even if the tradable sector is composed of a random group of commodities.

For the law of one price to be extended from tradables to all commodities, then, one must demonstrate that the tradable/nontradable distinction made in Sect. 6.1.5 has firm empirical foundation. Prior investigations have provided such support via the following results. First, based on input–output data, the ratios of both imports and exports to domestic sales are substantially higher for the tradable than the nontradable sector. Second, cross-country correlations of inflation rates or price indexes are higher for tradables than nontradables. Third, import price indexes are more highly correlated with price indexes of tradables than of nontradables. Fourth, as explanatory variables in a formal model of import demand, the price index of nontradables is uniformly nonsignificant whereas that of tradables is significant in a majority of cases. (See Goldstein and Officer 1979, pp. 421–422; Goldstein et al. 1980, pp. 193–196.).

To supplement these previous investigations, a test based specifically on the model of the present study is appropriate. Perhaps the most serious deficiency of the adopted (or indeed any) tradable/nontradable dichotomy is that the level of aggregation of existing data may be too high to permit a clear classification of industries into one sector or the other. In allocating a full industry to the category in which the preponderance of its sub-industries belongs, some of its output inevitably becomes included in the wrong sector. The most obvious example is services, which, while allocated totally to the nontradable sector, clearly have a tradable component, as the balance-of-payments table for any country shows.

Letting the export of services as a percentage of the production of nontradables represent the "tradable component" of the nontradable sector, are the prediction errors of Eq. (6.5) correlated with this component? If so, then the tradable/nontradable distinction is incorrect and Eq. (6.5) misspecified, calling into question the favorable findings for the law of one price based on this equation and reported in Sects. 6.1.6 and 6.1.7.

The following regression equation was estimated cross-sectionally over seventeen domestic countries (with Hungary, Poland, and Iran excluded due to data limitations):

$$\underset{}{Y} = \underset{(1.78)}{0.08} + \underset{(1.11)}{0.0056\mathrm{X}} \quad \overline{R}^2 = 0.01,$$

where

Υ = absolute value of residuals from Eq. (6.5) estimated as Equation number 1, listed in column 5 of Table 6.1
X = percentage of nontradables output that is exported, average of base year and 1975.

If services are illegitimately excluded from the tradable sector, then the slope of the regression should be significantly positive. Though positive, the slope is nonsignificant and the explanatory power of the regression is poor. The evidence, therefore, is that services are properly classified as nontradables, given a tradable/nontradable dichotomy.

6.1.9 Concluding Comments

Conventional testing of the law of one price involves considerable disaggregation of commodities, and modelling not going beyond linear regression in sophistication. Results of this approach are decidedly unfavorable to the law of one price once products have any element of differentiation.

In this paper a novel technique was introduced. The law of one price was tested using a breakdown of commodities at an aggregation no lower than that of tradables versus nontradables. Using the tradable/nontradable dichotomy, a model was developed that tested the law both domestically (that the prices of tradables and nontradables are equal) and internationally (that the price of tradables is identical across countries). In both cases the law of one price received strong support.

The positive results of the aggregative approach of this paper contrast with the negative findings of the disaggregative technique. While each approach has its own limitations, no study has found the weaknesses of the disaggregative framework to be unimportant, and there is reason to believe that the biases of this framework are unfavorable to the law of one price. In contrast, the present study demonstrates that the limitations of the aggregative approach do not affect its positive findings for the law of one price.

Appendix: The Data

1. Purchasing Power Parities—The only measures accepted were those based on direct price comparisons or on extrapolations of such measures via detailed extrapolations of components of GDP (rather than simple extrapolation via GDP deflators). Data

sources (by year) are as follows. 1975: Kravis and others (1982); 1950 (except Canada) and 1955: Gilbert and associates (1958); 1970: Kravis and others (1978); 1967:Kravis and others (1975); 1965 (Poland): Wiles (1971); 1950 and 1965 (Canada):Walters (1968); 1968 (Brazil, Mexico, Uruguay): Salazar-Carrillo (1978), with Colombia as base country, converted to United States as base via Colombia/U.S. 1968/1970 GDP deflator applied to Colombia/U.S. 1970 PPP; 1963 (Thailand): Usher (1968), with United Kingdom as base country, converted to United States as base via U.K./U.S. 1963/1967 GDP deflator applied to U.K./U.S. PPP for 1967.

2. Exchange Rates—International Monetary Fund (1980 and 1982 Yearbooks), series *af* (inverse of *ah* for United Kingdom) preferred, otherwise series *rf*. Poland and Hungary: *Pick's Currency Yearbook*, various issues; Poland: "effective official rate;" Hungary: "capitalistic tourist/noncommercial rate" (daily average computed for 1975).
3. Tradable/Nontradable Data—OECD (various issues); United Nations (various issues). Where segments of series are non- comparable, the earlier segment is linked to the later using a conversion ratio calculated from the earliest year of overlap.
4. GDP Deflator—International Monetary Fund (*Supplement on Price Statistics, 1981*). Hungary and Poland: United Nations (various issues).
5. Exports—Goods-and-services combined: International Monetary-Fund (1980 Yearbook), line 90c. Italy: 1950 obtained from OECD, *Statistics of National Accounts 1950–1961* on basis of overlap. Goods: International Monetary Fund (1980 Yearbook), line 70. Uruguay: line 70, d—in dollars, converted to domestic currency using exchange rate *rf*. India: International Monetary Fund (March 1971 and April 1979)—to obtain data for year beginning April 1, consistent with PPP, tradable/nontradable, and goods and services exports data. Services: Obtained by subtraction.

Notes

1. The thirteen studies are Bordo and Choudhri (1976), Curtis (1971), Dunn (1970, 1973), Isard (1977a, b), Kravis and Lipsey (1971, 1977, 1978), Norman (1975), Ormerod (1980), Richardson (1978), and Ripley (1974).

2. A similar point, with the exclusion of information costs, is made by Crouhy-Veyrac and others (1982).
3. It can be argued that a unit-factor-cost concept is the most appropriate methodology for PPP (Houthakker 1962, pp. 293–294). Now, under certain assumptions, a unit-factor-cost concept of PPP is equivalent to a PPP based on price levels that are a production-weighted average of commodity prices in each country (Houthakker 1962, p. 296; Officer 1974, pp. 871–872; 1976a, pp. 11–12; 1978, p. 564).
4. Of course, error terms are best incorporated structurally, i.e., included in Eqs. (6.3) and (6.4) and thence in Eq. (6.5). This pure procedure is not followed here for mathematical simplicity and because the estimation technique adopted is not thereby affected (see Sect. 6.1.6). Note also that for the parameter β to be estimable from the data, all observations must have the same current period, though their base periods can differ.
5. These industries are electricity, gas, and water; construction; wholesale and retail trade; restaurants and hotels; transportation, storage, and communication; finance, insurance, real estate, and business services; government services; and other producers of services.
6. For a thorough development of this measure, a discussion of its limitations, and empirical testing of the appropriateness of the $\hat{P}T$, $\hat{P}N$ series in general and the tradable/nontradable industry breakdown in particular, see Goldstein and Officer (1979) and Goldstein, Khan, and Officer (1980). The present study is not the first in which the $\hat{P}T$ and $\hat{P}N$ variables are elements in econometric testing (as distinct from being the subject of such testing). Predecessors are Officer (1976b), Goldstein and Officer (1979), Goldstein, Khan, and Officer (1980), and Stone (1982).
7. See Judge and others (1980, pp. 725–727). For curve-fitting purposes, it is acceptable to define the "error" simply as the difference between the left-hand and right-hand sides of Eq. (6.5), thereby justifying nonlinear least-squares. If one wished to obtain the additional, maximum-likelihood, properties of asymptotic unbiasedness, asymptotic efficiency, and sufficiency, then not only must a specific distribution of any error term be imposed but also the errors must enter Eq. (6.5) via Eqs. (6.3) and (6.4) structurally. The maximum-likelihood estimate would become quite complex.
8. For Canada, several years of PPP estimates are available, and the "base" and "current" years are chosen so as to maximize the intervening time period.

References

Bordo, M.D., and E. Choudhri. "The Behaviour of the Prices of Traded and Nontraded Goods: The Canadian Case 1962–74." Paper presented at the Southern Economic Association Annual Meetings, Atlanta, October 19, 1976.

Crouhy-Veyrac, L., Crouhy, M., and J. Mélitz. "More About the Law of One Price." *European Economic Review* 18 (July 1982): 325–344.
Curtis, J.M. "Direct Foreign Influences on Canadian Prices and Wages." In *Capital Transfers and Economic Policy: Canada, 1951–1962*. R.E. Caves and G.L. Reuber, eds. Cambridge, MA: Harvard University Press, 1971, pp. 196–231.
Dunn, R.M., Jr. "Flexible Exchange Rates and Oligopoly Pricing: A Study of Canadian Markets." *Journal of Political Economy* 78 (January–February 1970): 140–151.
______. "Flexible Exchange Rates and Traded Goods Prices: The Role of Oligopoly Pricing in the Canadian Experience." In *The Economics of Common Currencies*. H.G. Johnson and A.K. Swoboda, eds. Cambridge, MA: Harvard University Press, 1973, pp. 259–280.
Genberg, H. *World Inflation and the Small Open Economy*. Stockholm: Swedish Industrial Publications, 1975.
Gilbert, M., and associates. *Comparative National Products and Price Levels*. Paris: Organization for European Economic Cooperation, 1958.
Goldstein, M., and L.H. Officer. "New Measures of Price and Productivity for Tradable and Nontradable Goods." *Review of Income and Wealth* 25 (December 1979): 413–427.
______. M.S. Khan, and L.H. Officer. "Prices of Tradable and Nontradable Goods in the Demand for Total Imports." *Review of Economics and Statistics* 62 (May 1980): 190–199.
Houthakker, H.S. "Exchange Rate Adjustment." In *Factors Affecting the United States Balance of Payments*. Compilation of studies prepared for the Subcommittee on International Exchange and Payments, Joint Economic Committee, 87th Congress, 2nd session. Washington, DC: U.S. Government Printing Office, 1962, pp. 287–304.
International Monetary Fund. *International Financial Statistics*. Various issues.
Isard, P. "The Price Effects of Exchange-Rate Changes." In *The Effects of Exchange Rate Adjustments*. P.B. Clark, Logue, D.E., and R.J. Sweeney, eds. Washington, DC: U.S. Government Printing Office, 1977a, pp. 369–386.
______. "How Far Can We Push the 'Law of One Price'?" *American Economic Review* 67 (December 1977b): 942–948.
Judge, G.G., Griffiths, W.E., Hill, R.C. and T.C. Lee. *The Theory and Practice of Econometrics*. New York: Wiley, 1980.
Kravis, I.B., Kenessey, Z., Heston, A. and R. Summers. *A System of International Comparisons of Gross Product and Purchasing Power*. Baltimore and London: Johns Hopkins University Press, 1975.
______. Heston, A. and R. Summers. *International Comparisons of Real Product and Purchasing Power*. Baltimore and London: Johns Hopkins University Press, 1978.

______. Heston, A. and R. Summers. *World Product and Income*. Baltimore and London: Johns Hopkins University Press, 1982.
______. and R.E. Lipsey. *Price Competitiveness in World Trade*. New York: Columbia University Press, 1971.
______. and R.E. Lipsey. "Export Prices and the Transmission of Inflation." *American Economic Review* 67 (February 1977): 155–163.
______. and R.E. Lipsey. "Price Behavior in the Light of Balance of Payments Theories." *Journal of International Economics* 8 (May 1978): 193–246.
Norman, N.R. "On the Relationship Between Prices of Home-Produced and Foreign Commodities." *Oxford Economic Papers* 27 (November 1975): 426–439.
OECD. *National Accounts Statistics*. Various issues.
______. *Statistics of National Accounts 1950–1961.*
Officer, L.H. "Purchasing Power Parity and Factor Price Equalization." *Kyklos* 27 (Fasc. 4, 1974): 868–878.
______. "The Purchasing-Power-Parity Theory of Exchange Rates: A Review Article." International Monetary Fund *Staff Papers* 23 (March 1976a): 1–60.
______. "The Productivity Bias in Purchasing Power Parity: An Empirical Investigation." *International Monetary Fund Staff Papers* 23 (November 1976b): 545–557.
______. "The Relationship Between Absolute and Relative Purchasing Power Parity." *Review of Economics and Statistics* 60 (November 1978): 562–568.
Ormerod, P. "Manufactured Export Prices in the United Kingdom and the 'Law of One Price.'" *Manchester School of Economic and Social Studies* 48 (September 1980): 265–283.
Pick's Currency Yearbook. Various issues.
Richardson, J.D. "Some Empirical Evidence on Commodity Arbitrage and the Law of One Price." *Journal of International Economics* 8 (May 1978): 341–351.
Ripley, D. *The Movement of Export and Domestic Prices in the United States, Japan and Germany*. DM/74/66, International Monetary Fund, July 10, 1974.
Rosenberg, L.C. "Impact of the Smithsonian and February 1973 Devaluations on Imports: A Case Study of Steel." In *The Effects of Exchange Rate Adjustments*. P.B. Clark, Logue, D.E. and R.J. Sweeney, eds. Washington, DC: U.S. Government Printing Office, 1977, pp. 419–434.
Salazar-Carrillo, J. *Prices and Purchasing Power Parities in Latin America 1960-1972*. Washington, DC: Organization of American States, 1978.
Stone, J.A. "Nontraded Goods, Uncertainty, and Equilibration in the United States' Economy." *Journal of Macroeconomics* 4 (Summer 1982): 309–323.
United Nations. *Yearbook of National Accounts Statistics*. Various issues.

Usher, D. *The Price Mechanism and the Meaning of National Income Statistics*. Oxford: Oxford University Press, 1968.
Walters, D. *Canadian Income Levels and Growth: An International Perspective*. Ottawa: Queen's Printer, 1968.
Wiles, P.J.D. *The Prediction of Communist Economic Performance*. Cambridge: Cambridge University Press, 1971.

CHAPTER 7

National Price Level

7.1 The National Price Level: Theory and Estimation

Republished with permission of Elsevier, from *Journal of Macroeconomics*, Vol. 11, No. 3 (Summer 1989), pp. 351–373; permission conveyed through Copyright Clearance Center, Inc.

The national price level—defined as the ratio of the domestic country's price level expressed in base-country currency to the base-country's price level, or, equivalently, the ratio of purchasing power parity (PPP) to the exchange rate—has developed a literature of its own in the 1980s. Work has flourished in every aspect: (a) data generation (principally, the International Comparison Project [ICP], with Phase 3 published as Kravis et al. [1982], and Phases 4 and 5 continuing at the United Nations); (b) use of the national price level in short-cut estimation of real income for international comparison (for example, Isenman 1980; Clague 1986b); (c) theoretical analysis (especially Kravis and Lipsey 1983; Bhagwati 1984; Clague 1985, 1986c); and (d) econometric explanation (for example, Salazar-Carrillo 1982b; Clague 1986a, 1988a, b; Kravis and Lipsey 1987).

This paper shows that the existing literature suffers uniformly from methodological and econometric problems, and attempts to correct these defects. The approach is to take seriously the assumptions and

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_7

implications of the accepted theory of the national price level, contrary to the practice of practitioners in the field. The state of the art is improved by deriving the analytical relationship between the national price level and the nontradable/tradable price-level ratio as it varies with the specific PPP index-number concept, and then making use of this relationship in econometric estimation of the national price level.

7.1.1 Conventional Approach to the National Price Level

Consider the following notation:

PL = national price level, ratio of price level of domestic country expressed in base-country currency to price level of base country, dimensionless
PPP = purchasing power parity for gross domestic product (GDP), number of units of domestic currency per unit of base-country currency
R = exchange rate, number of units of domestic currency per unit of base-country currency
Y_{jk} = GDP of country j valued at prices of country k;
b = base country
d = domestic country.

Then, by definition,

$$\mathrm{PL} = \mathrm{PPP}/R. \tag{7.1}$$

The term "national price level" emanates from the equivalent definition of PL as the ratio of "nominal" to "real" income, that is, the ratio of GDP converted to base-country currency via the exchange rate (sometimes called "conventional income") to GDP converted to base-country currency via PPP ("real income"):

$$\mathrm{PL} = (Y_{dd}/R)/Y_{db},$$

where

$$Y_{db} = Y_{dd}/\mathrm{PPP}. \tag{7.2}$$

The conventional approach to explaining the national price level follows a three-step process, but the literature does not formalize the procedure. Three propositions are involved. First, all commodities are divided into two sharply defined classes: tradables and nontradables (sometimes loosely designated as goods and services, respectively).[1] Therefore

$$Y_{jj} = \mathrm{YT}_j + \mathrm{YN}_j, \quad j = b, d; \tag{7.3}$$

where YT_j (YN_j) denotes the output of tradables (nontradables) of country j valued at prices of country j.

Second, the "law of one price" for tradables is imposed, that is, the price level of tradables expressed in a common currency is equalized in the domestic and base countries.[2] Letting PT (PN) denote the ratio of the price level of tradables (nontradables) in the domestic country to that in the base country, the law of one price for tradables is

$$\mathrm{PT} = R. \tag{7.4}$$

Third, the principal influences on the national price level are deemed to be long run or structural, and to operate through affecting the ratio of the nontradable to the tradable price level in the domestic country relative to the base country, that is, the nontradable/tradable price-level ratio, P, defined as

$$P = \mathrm{PN}/\mathrm{PT}. \tag{7.5}$$

Because the nontradable/tradable price-level ratio, P, and the national price level, PL, are positively related, the explanatory variables for P are ultimate determinants of PL with the same directional effect.[3]

Accepting propositions one and two, as well as proposition three to the extent it is consistent with the first two, one can show that the literature suffers from a variety of limitations:

1. The third proposition of the approach (that long-term determinants of the national price level operate through the nontradable/tradable price-level ratio) does not necessarily follow from the second proposition (the law of one price for tradables). Indeed, the third proposition does not always apply.

2. Short-run (monetary) determinants of the national price level are considered in both theory and econometric testing. Yet such variables can enter the analysis only tenuously compared to long-run (structural) variables.
3. While hypothesis specification for long-run explanatory variables directly concerns the impact of these variables on the nontradable/tradable price-level ratio (P), econometric estimation invariably uses the national price level (PL) as the dependent variable. True, PL is the variable of interest; however, direct hypothesis testing requires an equation in which P is the dependent variable.[4]
4. There is an analytical relationship between P and PL that depends on the specific index selected for PPP. This relationship is not derived or even mentioned in the literature, and therefore it is ignored—and sometimes contradicted—in econometric estimation of the national price level.
5. The analytical relationship between P and PL has implications for hypothesis specification for PL beyond the explanatory variables operating through P. Consistent with point one, the relationship can stipulate long-run explanatory variables for a PL regression quite apart from the P route, a phenomenon alien to the existing literature.
6. The fact that there is an analytical expression for PL in terms of P and (depending on the index selected) other variables implies that conventional econometric estimation of PL, which confines itself to linear or log-linear functional form, is generally incorrect, as the PL-P analytical expression is inherently nonlinear for most PPP measures.
7. Existing empirical studies make no effort to delineate the relative importance of explanatory variables for PL. Rather, it is *assumed* that real per capita income is the variable with greatest impact on PL.

How could deficiencies of this magnitude exist in an established body of literature? The reason is twofold. First, practitioners in the field have not applied index-number algebra to their body of theory in a general way, even though purchasing power parity (by definition, the numerator of the national price level) is inherently an index number. Second, practitioners do not take seriously the three propositions that they explicitly adopt as the basis for theorizing about the national price level, and therefore have not carried the implications of these propositions to their logical conclusions.

7.1.2 The Various PPP Indexes

Let ST_j (SN_j) denote the share of tradables (nontradables) in the output of country *j*, defined as follows:

$$\mathrm{ST}_j = \mathrm{YT}_j / Y_{jj}, \quad j = b, d; \tag{7.6}$$

$$\mathrm{SN}_j = \mathrm{YN}_j / Y_{\mathrm{jj}}, \quad j = b, d; \tag{7.7}$$

From (7.3), it follows that

$$\mathrm{ST}_j = 1 - \mathrm{SN}_j, \quad j = b, d. \tag{7.8}$$

Applying proposition one of the conventional approach (a tradable/nontradable dichotomy of output and therefore price) and a two-country model (the base country and only one domestic country), well-known PPP indexes are the Laspeyres, Paasche, Fisher ideal, Walsh, and Geary-Khamis, with the following definitions.[5]

$$\text{Laspeyres: PPP} = \mathrm{ST}_b \cdot \mathrm{PT} + \mathrm{SN}_b \cdot \mathrm{PN}. \tag{7.9}$$

$$\text{Paasche: PPP} = \frac{1}{\mathrm{ST}_d/\mathrm{PT} + \mathrm{SN}_d/\mathrm{PN}}. \tag{7.10}$$

$$\text{Fisher: PPP} = \left(\frac{\mathrm{ST}_b \cdot \mathrm{PT} + \mathrm{SN}_b \cdot \mathrm{PN}}{\mathrm{ST}_d/\mathrm{PT} + \mathrm{SN}_d/\mathrm{PN}}\right)^{1/2}. \tag{7.11}$$

The variables PPP, PT, and PN are specific to each index. The Laspeyres PPP (and PT, PN) is the sum of domestic-country/base-country price relatives, each weighted by the base-country output share; while the Paasche PPP (and PT, PN) is the inverse of the sum of base-country/domestic-country price relatives, each weighted by the domestic-country output share. The Fisher ideal index is the geometric mean of the Laspeyres and Paasche indexes.

The general form of the Walsh index is

$$\text{Walsh: PPP} = \mathrm{PT}^{\mathrm{WT}} \cdot \mathrm{PN}^{\mathrm{WN}}, \tag{7.12}$$

where

$$\mathrm{WT} = 1 - \mathrm{WN}. \tag{7.13}$$

The Walsh index with arithmetic-mean weights is the product of domestic-country/base-country price relatives, each weighted exponentially by the arithmetic mean of the countries' output shares:

$$\mathrm{WT} = (\mathrm{ST}_b + \mathrm{ST}_d)/2,$$

$$\mathrm{WN} = (\mathrm{SN}_b + \mathrm{SN}_d)/2.$$

The Walsh index with geometric-mean weights is the product of domestic-country/base-country price relatives, each weighted exponentially by the geometric mean of the countries' output shares; the means are normalized so that (7.13) is satisfied:

$$\mathrm{WT} = \frac{(\mathrm{ST}_b \cdot \mathrm{ST}_d)^{1/2}}{(\mathrm{ST}_b \cdot \mathrm{ST}_d)^{1/2} + (\mathrm{SN}_b \cdot \mathrm{SN}_d)^{1/2}},$$

$$\mathrm{WN} = \frac{(\mathrm{SN}_b \cdot \mathrm{SN}_d)^{1/2}}{(\mathrm{ST}_b \cdot \mathrm{ST}_d)^{1/2} + (\mathrm{SN}_b \cdot \mathrm{SN}_d)^{1/2}},$$

$$\text{Geary-Khamis: } \mathrm{PPP} = \frac{1}{\mathrm{ST}_d \cdot (\mathrm{PT}_I/\mathrm{PT}) + \mathrm{SN}_d \cdot (\mathrm{PN}_I/\mathrm{PN})}, \tag{7.14}$$

where $\mathrm{PT}_I(\mathrm{PN}_I)$ is the ratio of international price level of tradables (nontradables) to price level of tradables (nontradables) in the base country.

The Geary-Khamis PPP is a Paasche index with the base country replaced by the (two-country) world, as $\mathrm{PT}_I/\mathrm{PT}$ and $\mathrm{PN}_I/\mathrm{PN}$ are international/domestic-country price relatives. This index is expressed, along with the other PPP indexes, as the number of units of domestic currency per unit of base-country currency; because PPP for the base country (the base-country/international PPP) is identically equal to unity.[6]

$$\mathrm{PT}_I = \frac{\mathrm{YT}_d/\mathrm{PPP} + \mathrm{YT}_b}{\mathrm{YT}_d/\mathrm{PT} + \mathrm{YT}_b} \tag{7.15}$$

$$\mathrm{PN}_I = \frac{\mathrm{YN}_d/\mathrm{PPP} + \mathrm{YN}_b}{\mathrm{YN}_d/\mathrm{PN} + \mathrm{YN}_b} \tag{7.16}$$

The international/base-country price level of tradables (nontradables) is the ratio of the sum of domestic-country and base-country output at international prices to the sum at base-country prices. There is an asymmetry in that there is only one international-currency/domestic-currency conversion factor (1/PPP) and it pertains to all output (as does the international-currency/base-currency conversion factor, unity), whereas the base-currency/domestic-currency conversion factor (1/PT or 1/PN) is specific to tradables or nontradables.

Equations (7.14), (7.15), and (7.16) constitute a system to be solved simultaneously for PPP, PT_I, and PN_I. The variable that is of interest, PPP, is the one positive root of the following quadratic:

$$\mathrm{A}\ \cdot PPP^2 + \mathrm{B} \cdot PPP + C = 0; \tag{7.17}$$

where

$$\begin{aligned} A &= \mathrm{ST}_d \cdot \mathrm{SN}_d \cdot Y/\mathrm{PT}, \\ B &= \mathrm{ST}_d \cdot \mathrm{ST}_b \cdot \mathrm{SN}_b \cdot P + \mathrm{ST}_b \cdot \mathrm{SN}_d \cdot \mathrm{SN}_b \\ &\quad - \mathrm{ST}_d \cdot \mathrm{SN}_d \cdot \mathrm{SN}_b \cdot Y \cdot P - \mathrm{ST}_d \cdot \mathrm{ST}_b \cdot \mathrm{SN}_d \cdot Y, \\ C &= -\mathrm{ST}_b \cdot \mathrm{SN}_b \cdot \mathrm{PN}, \\ Y &= Y_{db}/Y_{bb}. \end{aligned}$$

7.1.3 *Issues of Theory and Estimation*

First issue: What is the analytical expression for the national price level, PL, in terms of the nontradable/tradable price-level ratio, P, assuming the law of one price for tradables?

For the Laspeyres, Paasche, Fisher, and Walsh indexes, the answer is obtained by taking Eq. (7.1), PL = PPP/R, substituting the right-hand sides of (7.9)–(7.12) for the respective indexes, applying (7.8) to eliminate ST_b and/or ST_d and (7.13) to eliminate WT (for the Walsh index), simplifying the resulting expression, and substituting P for PN/PT (via [7.5]) and PT for R (via [7.4]). The results are

$$\text{Laspeyres: } \mathrm{PL} = 1 + \mathrm{SN}_b(P - 1), \tag{7.18}$$

$$\text{Paasche: } \mathrm{PL} = \frac{P}{P(1 - \mathrm{SN}_d) + SN_d}, \tag{7.19}$$

$$\text{Fisher: PL} = \left[\frac{\text{SN}_b(P-1)+1}{\text{SN}_d(1/P-1)+1}\right]^{1/2}, \tag{7.20}$$

$$\text{Walsh: PL} = P^{\text{WN}}; \tag{7.21}$$

where PL and P are specific to the index under consideration.

The Laspeyres, Paasche, and Fisher PPP indexes possess the property of "characteristicity," that is, the price comparison between a given domestic country and the base country is determined solely by data in these two countries. Therefore, for these measures, moving to an N-country sample [$(N-1)$ domestic countries]—discussion of which is crucial for econometric analysis—does not affect the formula for PL in terms of P. For the Walsh index, "characteristicity" is not present. With $(N-1)$ domestic countries, WT and WN are redefined so that the arithmetic or geometric mean of tradable (or nontradable) shares runs over all N countries rather than only the base country and the domestic country. Therefore, the Walsh formula for PL retains its form with N countries; only WN is redefined.

As for the Geary-Khamis index, solving for the positive root of (7.17) and applying (7.4) yields:

Geary-Khamis:

$$\text{PL} = \frac{-D + \left(D^2 + 4\cdot\text{ST}_b\cdot\text{ST}_d\cdot\text{SN}_b\cdot\text{SN}_d\cdot P\cdot Y\right)^{1/2}}{2\cdot\text{ST}_d\cdot\text{SN}_d\cdot Y}, \tag{7.22}$$

where

$$\begin{aligned} D = {} & \text{ST}_b\cdot\text{ST}_d\cdot\text{SN}_b\cdot P + \text{ST}_b\cdot\text{SN}_b\cdot\text{SN}_d \\ & - \text{ST}_d\cdot\text{SN}_b\cdot\text{SN}_d\cdot Y\cdot P - \text{ST}_b\cdot\text{ST}_d\cdot\text{SN}_d\cdot Y. \end{aligned}$$

It is noted that P, SN_b and SN_d, and Y enter the formula for PL (with ST_b and ST_d eliminable via [7.8]). With N domestic countries (the variables for which may be subscripted by d), Eqs. (7.15) and (7.16) become

$$\text{PT}_I = \frac{\sum_d \text{YT}_d/\text{PPP}_d + \text{YT}_b}{\sum_d \text{YT}_d/\text{PT}_d + \text{YT}_b},$$

$$\mathrm{PN}_I = \frac{\sum_d \mathrm{YN}_d/\mathrm{PPP}_d + \mathrm{YN}_b}{\sum_d \mathrm{YN}_d/\mathrm{PN}_d + \mathrm{YN}_b};$$

and there are $(N - 1)$ equations of type (7.14) each specific to PPP_d, ST_d, SN_d, PT_d, and PN_d. Substituting for PT_I and PN_I leaves a system of $(N - 1)$ nonlinear equations determining the PPP_d, $d = 1, \ldots, N - 1$, the analytical solution of which is not attempted. Presumably, PL for a given domestic country (PL_d) depends on *all* the P_d, SN_d, and Υ_d, as well as on SN_b. The Geary-Khamis index exhibits extreme "noncharacteristicity."

Pleasingly, Eqs. (7.18)–(7.22) show that, for all the PPP indexes, $dPL/dP > 0$, justifying the use of the variables determining P as explanatory variables for PL.

Second issue: Assuming the law of one price for tradables, is PL a function of P alone so that it is legitimate to hypothesize that all (structural) variables determine PL by affecting P?

To answer this question, an N-country sample is assumed at the outset. For the Laspeyres, Paasche, and Fisher indexes, this sample serves merely to separate variables (PL, P, and SN_d) from parameters (SN_b), where a parameter is invariant over the $N - 1$ domestic countries. For the Walsh index, WN plays the role of SN_b, invariant over countries, and so may be regarded as a parameter. As Eqs. (7.18)–(7.21) show, then only for the Laspeyres and Walsh indexes is PL a function of the sole variable P. In the case of the Paasche and Fisher measures, PL depends on two variables, P and SN_d. Therefore SN_d, the share of nontradables in the domestic country's output, must enter a PL regression in its own right, that is, independent of whether it is a determinant of P.

For the Geary-Khamis index, Eq. (7.22) shows that for the two-country case, PL is a function not only of P but also of SN_d and Y. Extending the sample to N countries (meaning N observations) results in PL depending on the $3 \cdot (N - 1)$ variables, P_d, SN_d, and Υ_d, $d = 1, \ldots, N - 1$, a highly negative degrees-of-freedom situation.

Third issue: Given that conventional econometric work involves PL as the dependent variable and assuming that P can be expressed either as a linear or log-linear combination of its explanatory variables, is the conventional specification of a linear or log-linear functional form for PL (regressed on the determinants of P) correct?

With SN_b a parameter, a linear regression for the Laspeyres index is correct (that is, would provide proper tests of significance), as Eq. (7.18) shows. However, a log-linear formulation would be incorrect; for the logarithm of the right-hand side of (7.18) is nonlinearly related to log$P$. In contrast, a linear form for the Walsh index is illegitimate, with PL and P nonlinearly related [see Eq. (7.21)]. Yet a log-linear form would be correct; for logPL = WN · log$P$, with WN a parameter.

However, the formulas for PL using the Paasche and Fisher indexes [Eqs. (7.19) and (7.20)] are inherently nonlinear in their right-hand-side variables, P and SN_d, so neither a linear nor a log-linear regression of PL on the determinants of P is consistent with the analytical relationship between PL and P. As for the Geary-Khamis index, the analytical expression for PL even in the two-country case [Eq. (7.22)] is too complex for consistency with a linear or log-linear formulation. For an N-country sample, the expression for PL resulting from solving the $(N - 1)$ equation system described previously would be even more inherently nonlinear (to say nothing of the degrees-of-freedom problem).[7]

Fourth issue: Should short-run variables be included in the determination of P?

The only such variable suggested in the literature is an international transfer, measured by a country's current-account imbalance; Salter (1959) and Clague (1986c) present models that clearly provide scope for a transfer to affect P. Alternative models, however, exhibit no such obvious impact of a transfer on P.[8] More fundamentally, a *cross-sectional* econometric determination of P (the sole interest in the literature) warrants stable, country-characteristic explanatory variables that are structural and long term in nature; and it is with reason that variables of this ilk dominate in empirical work.

Fifth issue: Granted that P may be presumed to depend only on long-run (or structural) variables, is it nevertheless legitimate to include short-run (or monetary) variables in an equation in which PL is the dependent variable?

Kravis and Lipsey (1983) and Clague (1986a, 1988a) take the view that short-run variables can enter simply as additional terms in a regression equation for PL that otherwise include long-run influences on PL operating through P. These authors use Dornbusch's (1976) overshooting model to justify inclusion of money-supply growth as an explanatory variable in the PL regression equation.[9]

However, propositions one and two of the accepted doctrine on the national price level negate the existence of short-run determinants of PL except insofar as they affect P. Equations (7.18)–(7.21) and an N-country version of (7.22) apply, and PL depends only on long-run variables: P (or those variables that determine P), SN_d, Υ, and the structural parameter SN_b.[10] There is no scope for short-run variables to influence PL except insofar as they affect P directly; and, for cross-sectional analysis, short-run explanatory variables for P are to be avoided (see *fourth issue* above).

Suppose, now, that the law of one price for tradables, proposition two, is abandoned (though a *tendency* toward this law must be assumed; else the distinction between tradables and nontradables, proposition one, is eliminated by default). Then the right-hand sides of Eqs. (7.18)–(7.22), and presumably the N-country analogue of (7.22), now include the multiplicative factor (PT/R); whereas previously this term was eliminated by Eq. (7.4). The additional variables that would enter the equation for PL must act to explain a nonunity PT/R, that is, nonfulfillment of the law of one price. Such factors as monopoly and oligopoly, transport and other transactions costs, trade restrictions, product differentiation, and differing commodity compositions of the price of tradables across countries—all of which are structural—act to explain the failure of the law of one price.[11] Again there is no scope for monetary variables except insofar as they affect P directly, independent of whether the law of one price holds.

Sixth issue: In the existing literature, long-run determinants of the national price level operate through P, that is, they are really theories about P. Therefore the question arises: Again assuming the law of one price for tradables, how can such hypothesis testing be more precise if PL, but not P, data are available?

Equations (7.18)–(7.21) re-solved for P yield

$$\text{Laspeyres: } P = \frac{\mathrm{PL}}{\mathrm{SN}_b} + \frac{\mathrm{SN}_b - 1}{\mathrm{SN}_b}; \tag{7.23}$$

$$\text{Paasche: } P = \frac{PL \cdot \mathrm{SN}_d}{\mathrm{PL}(\mathrm{SN}_d - 1) + 1}; \tag{7.24}$$

$$\text{Fisher: } P = \frac{E + \left(E^2 + 4 \cdot \mathrm{SN}_b \cdot \mathrm{SN}_d \cdot \mathrm{PL}^2\right)^{1/2}}{2 \cdot \mathrm{SN}_b} \tag{7.25}$$

where $E = \mathrm{PL}^2(1 - \mathrm{SN}_d) + \mathrm{SN}_b - 1$;

$$\text{Walsh: } P = \mathrm{PL}^{1/\mathrm{WN}}. \tag{7.26}$$

Equations (7.23)–(7.26) enable conversion of PL to P, which can then be used as the dependent variable in regressions. For the Geary-Khamis index, the relationship between PL_d for a given domestic country (d) and $P_d, d = 1, \ldots, N-1$, which would involve solving $(N - 1)$ nonlinear equations, is not known. There are two cases in which tests of significance are the same whether PL or P is the dependent variable: a linear equation with the Laspeyres index or a log-linear formulation with the Walsh index (see *third issue* above).

Seventh issue: Is it appropriate to adopt the Geary-Khamis index-number concept as the data basis for estimation and use of the national price level?

Geary-Khamis PPP and real-income data are exclusively employed in the most recent econometric studies of the national price level, whether of its estimation (Kravis and Lipsey 1983, 1987; Clague 1986a, 1988a) or of its use in facilitating short-cut estimates of real income (Clague 1986b). This is not surprising, as the ICP, having adopted Geary-Khamis as its official index-number concept in Phase 1 (Kravis et al. 1975), reached the point that by Phase 4 (United Nations 1986), even partial data on another index-number basis were not presented.

A disadvantage of Geary-Khamis is its mathematical complexity for $N > 2$. Kravis et al. (1982, 93) properly point out that sufficient resources overcome the *computational* complexity of Geary-Khamis. For *theoretical* purposes, however, analytical solutions are desirable, and these are extremely difficult to obtain for Geary-Khamis relative to the other indexes considered. For example, derivations of the PL-P analytical relationship and its inverse were not attempted above for $N > 2$.

Certainly, it is required that there be theoretical and empirical consistency of a researcher's selected index-number concept. In this light, Clague's adoption of Geary-Khamis data is inappropriate, because the PPP concept that he incorporates in the theory underlying his empirical work is the Walsh index.[12] A Walsh PPP index demands Walsh data. To make his use of Geary-Khamis data legitimate, Clague would have to use a Geary-Khamis PPP concept in his underlying theory, and the mathematics involved would be difficult indeed!

Empirically, the Geary-Khamis index (along with the Walsh index) has the advantage of being base-country invariant, but it lacks "characteristicity," even more than the Walsh index, to the point that a *negative* degrees-of-freedom situation must be overridden in econometric work. Other criticisms have been levied against the ICP Geary-Khamis data, including Isenman's (1980, pp. 65–66) admonition that basing decisions on these data would result in a misallocation of resources for less-developed countries. Kravis (1984, pp. 33–35; 1986, pp. 21–23) replies to the Isenman and related critiques, but does not address the weaknesses of Geary-Khamis suggested here. Most seriously, it is impossible for a methodologically correct econometric procedure to fit Geary-Khamis data (see *eighth issue* below). Primarily for this reason, the Fisher ideal index, and not Geary-Khamis, is used in the empirical part of this study.

Eighth issue: What is a methodologically correct econometric procedure for the estimation of PL?

Existing authors adopt an ad hoc approach to the econometrics of PL. Without rationale, they simply regress PL on the determinants of *P* and on monetary (short-run) variables. Not only do they ignore the PL-P analytical relationship and the other findings of *issues one to seven* but also they assert, without testing, a relative importance of the determinants of PL. In contrast, a correct econometric procedure consists of the following operations:

Step 1: Select the PPP-index concept.
Step 2: Obtain the PL-P analytical relationship for that PPP index.
Step 3: Regress PL on P and the other variables entering the PL-P analytical relationship. The resulting equation is an econometric approximation to that relationship.
Step 4: List, and provide theoretical justification for, the variables explaining *P*.
Step 5: Regress *P* on these variables. This provides not only a check on their appropriateness as explanatory variables for PL but also an input into econometric estimation of PL based on the PL-P analytical relationship (Step 8).
Step 6: List the variables other than those in Step 4, that are present in the PL-P analytical relationship. They warrant inclusion as PL explanatory variables in addition to those specified in Step 4 (or as reinforcing the effects of variables entering the PL regression on both grounds).
Step 7: Regress PL on the explanatory variables emanating from Steps 4 and 6. The explanatory power of the estimated equation in Step 3 indicates the error in the functional form (say, linear or log-linear)

of this (Step 7) procedure relative to the form of the PL-P analytical relationship.

Step 8: Insert the estimated observation vector for P resulting from the Step 5 regression, the true observation vector of SN_d (if indicated), and the parameters SN_b or WN (as indicated) into the PL-P analytical relationship (Step 2).

Step 9: Compare the predictive accuracy of the alternative estimates of PL provided by Steps 7 and 8.

Step 10: Use beta coefficients to delineate the relative importance of the explanatory variables for PL.

With an unknown PL-P analytical relationship for the Geary-Khamis index, Steps 2, 8, and 9 are impossible to perform. Furthermore, the complicated nonlinearity of the relationship makes Step 3 of dubious utility, even if the degrees-of-freedom problem is solved by confining explanatory variables to those pertaining to the domestic country of observation, the procedure (though without rationale) followed by those authors that use Geary-Khamis data. Fortunately, Fisher-ideal-index data were published along with Geary- Khamis in Phase 3 of the ICP. Ironically, while ignored by previous researchers in favor of the Geary-Khamis index, the Fisher data meet fully the econometric methodology and so are adopted here. A sample of 31 countries for the year 1975 is used, consisting of 30 domestic countries and the base country, the United States.[13] All regressions are fit in log-linear form.

7.1.4 *Econometric Approximation of PL-P Analytical Relationship*

Equation (7.20) indicates that the PL-P analytical relationship for the Fisher index may be approximated by a PL regression in which P and SN_d are explanatory variables, with a positive and a negative effect on PL, respectively. Equation El in Table 7.1 presents this log-linear approximation, with t-values in parentheses adjacent to the estimated elasticities. Three measures of goodness-of-fit are computed: R^2, the conventional R-squared based on the transformed (logarithmic) variables and corrected for degrees of freedom; r, the correlation coefficient between the true, nontransformed dependent variable, PL, and the antilog of the estimated (est) dependent variable, exp[est(logPL)]; and U, Theil's inequality coefficient of the N country-subscripted pairs, PL and exp[est(logPL)].[14]

Table 7.1 Regression equations

Equation	*E1*	*E2*	*E3*
Dependent variable	logPL	log$P$	logPL
Constant	−0.23 (6.53)	−2.49 (2.17)	−1.42 (2.46)
Price ratio (log$P$)	0.51 (40.89)		
Per capita income (logYC)		0.38 (2.86)	0.17 (2.57)
Share of services (logSSER)		0.16 (2.12)	0.07 (1.93)
Natural resources (logNARE)		−0.33 (2.68)	−0.20 (3.17)
Literacy (logLIT)		−0.45 (2.50)	−0.22 (2.46)
Share of nontradables (logSN)	−0.32 (6.38)	1.83 (4.81)	0.59 (3.12)
R^2	0.99	0.92	0.89
r	0.99	0.95	0.95
U	0.03	0.11	0.11
Degrees of freedom	28	25	25

Equation El (see Table 7.1) provides excellent justification for the use of Fisher data in regressing PL on the variables determining P and on SN. Both P and SN have highly significant coefficients (beyond the one percent level) with the correct sign, and the goodness-of-fit is impressive on all three criteria.

7.1.5 Explanatory Variables for the Nontradable/Tradable Price-Level Ratio and the National Price Level

Per Capita Real Income (*YC*): There are two separate rationales for per capita real income. First, it proxies the ratio of productivity in tradables to productivity in nontradables (higher in high-income countries), which has a positive effect on P and, through it, PL ("the productivity-differential model," discussed in Kravis and Lipsey 1983, pp. 11–14). Second, with nontradables labor-intensive relative to tradables and high-income countries labor-scarce, labor is relatively expensive in these countries and P high (the factor-proportions explanation, outlined in Bhagwati 1984).

Share of International Services (SSER): The ratio of balance-of-payments-services credits (excluding investment income) to GDP has both a demand and supply effect on P. Services provided to foreigners constitute a greater demand for "nontradables," resulting in a higher P. However, SSER can just as well measure the supply or abundance of "tradable services" (part of nontradables) as the demand, hence a higher

SSER implies a lower *P*. Therefore, the directional effect of *SSER* is indeterminate.[15]

Natural Resources (NARE): The share of the production of natural-resource industries (agriculture, hunting, forestry, fishing, mining, and quarrying) in GDP similarly has an uncertain directional effect on *P*. With tradables more natural-resource intensive than nontradables, a greater abundance of natural resources implies a higher *P*. However, if NARE is interpreted as representing demand for the country's natural-resource-intensive commodities, the effect on *P* is reversed.[16]

Literacy (LIT): With services (nontradables) skilled-labor intensive, a higher-quality labor force (a clear supply-side variable) implies a lower *P*. This variable is used in the literature.

Share of Nontradables (SN): The ICP definition of nontradable output is adopted: final expenditure on services plus construction. Kravis and Lipsey (1983, p. 15) state that, with a low elasticity of substitution between tradables and nontradables (an empirical result), a high share of nontradables in output (SN) implies a higher *P*. Clague (1986a, p. 321) rejects this argument, because SN itself is determined by the other influences on *P* (though he writes PL). Though Clague is correct that SN and *P* are jointly-determined variables, the issue is one of simultaneity, which awaits general-equilibrium modelling for its solution. Until that time, SN may enter the *P* equation in full realization that single-equation specification and estimation are convenient abstractions.

Certainly, SN must always be present in the equation for PL itself, because of the analytical expression for PL, Eq. (7.20). This equation shows that the *direct* effect of SN (SN_d) on PL is negative. With an indirect positive impact on PL (that operating through *P*), the net impact of SN on PL (coefficient of SN on PL) should be algebraically smaller than the indirect effect (indicated by the coefficient of SN on *P*).

7.1.6 Alternative Estimators of the National Price Level

Equation E2, exhibiting the regression of *P* on its determinants, has an exceptionally high R^2 for cross-sectional data, and all five explanatory variables are significant at the five percent level. Equation E3, for PL, is also satisfactory, with R^2 almost as high, four coefficients significant at the five percent level, and the remaining coefficient (that for SSER) nearly so. Further, all coefficients in these equations have a theoretically correct sign. (The SSER and NARE coefficients have signs suggesting that these variables embody demand rather than supply.)

The PL-P analytical relationship for a given PPP index was honored in Sect. 7.1.4 in respect of inclusion of variables but not in respect of functional form, for a log-linear approximation was employed. Now the relationship can be given its full due as follows. Letting est(log$P$) be the estimated log$P$ observation vector from Equation E2, then exp[est(log$P$)] is plugged into Eq. (7.20) along with the true observation vector, SN, and the parameter SN_b (SN for the United States). The resulting estimate of PL may be compared with the true PL by means of their correlation coefficient (r) and inequality coefficient (U), shown in Equation E2.

An alternative, direct estimate of PL is exp[est(logPL)], where est(logPL) is the estimated observation vector on logPL obtained by applying Equation E3. This PL estimate, too, can be matched with the true PL via r and U. It turns out that these alternative estimates of PL, with identical goodness-of-fit statistics to two decimal places, are equally (and highly!) proficient at predicting the true PL.

7.1.7 Relative Importance of Determinants of the National Price Level

Previous authors have not investigated the relative importance of the determinants of the national price level. Rather, they *assume*, explicitly or implicitly, that real per-capita income is the most important explanatory variable.[17]

Beta coefficients provide an objective, if somewhat arbitrary, econometric answer to the question of relative contribution of individual variables to the explanation of the dependent variable. Because the estimated coefficients are elasticities rather than slopes, the beta coefficients are computed as the product of the estimated coefficient and the ratio of the coefficient of variation (ratio of standard deviation to mean) of the nontransformed (level rather than logarithmic) regressor under consideration to the coefficient of variation of the nontransformed dependent variable (PL).

Beta coefficients so computed are, in order of absolute value, natural resources (−0.35), share of nontradables (0.32), per capita income (0.32), literacy (−0.18), and share of services (0.11). The accepted view of per-capita income as the predominant explanatory variable receives no support. Natural resources and the share of nontradables are at least as important.

7.1.8 Principal Conclusions

a. There is an analytical relationship between the national price level, PL, and the nontradable/tradable price-level ratio, *P*, that varies with the PPP index adopted and that theoreticians and econometricians ignore at their peril.
b. Precise hypothesis testing for those determinants of PL that operate through *P* requires estimation of an equation in which *P* rather than PL is the dependent variable.
c. Not all long-run or structural influences on PL operate (or operate exclusively) through *P*; some instead (or also) affect PL directly.
d. Dropping the assumption of the law of one price for tradables alters the PL-P analytical relationship but does not in itself justify inclusion of short-run or monetary variables in the explanation of the national price level.
e. Conventional functional forms of the estimating equation for PL violate the PL-P analytical relationship and can lead to incorrect hypothesis testing. The amount of approximation involved in these functional forms is empirically testable.
f. There are two methods for estimating PL. Traditionally, PL is regressed directly on its explanatory variables. Alternatively, the regression for *P* is incorporated into the PL-P analytical relationship.
g. The conventional wisdom that per capita real income is the most important variable in the explanation of PL receives no support.
h. There exist theoretical and econometric problems in the use of the Geary-Khamis index to develop theories and test hypotheses concerning PL and *P*.

Appendix: The Data

All variables refer to the calendar year 1975; except for India and Iran, for which most data are for the year beginning April 1 or March 21, respectively.

National Price Level: PL = PPP/*R1*, where PPP is the "augmented" Fisher-ideal-index PPP for GDP and *R1* is the annual-average exchange rate, both measured as the number of units of domestic currency per dollar (Kravis et al. 1982, pp. 10, 253–85). The source also provides a list of the 31 countries in the sample.

Share of Nontradables: SN (Kravis et al. 1982, p. 194).

Nontradable/Tradable Price-Level Ratio: *P* is computed from Eq. (7.25) with SN_d = SN, and SN_b = SN for United States.

Per Capita Real Income: YC = (YD/PPP)/POP, where YD is GDP at national prices, millions of domestic-currency units, and POP is mid-year population, millions of persons (Kraviset al. 1982, pp. 10, 12).

Share of Services: SSER = *(R2* · SER)/YD, where *R2* is the annual-average exchange rate, number of units of domestic currency per Special Drawing Right (SDR), and SER is total service credits (excluding investment income) in the balance of payments, in SDRs (International Monetary Fund [1980] for Iran; International Monetary Fund [1983] for other countries). Unlike the other flow data, *SER* is on a calendar-year basis for India and Iran. For Belgium and Luxembourg, figures are consolidated and currencies are interchangeable; therefore, in the SSER formula, YD is the sum of the figures for the two countries and the resultant value of SSER is applied to both countries.

Natural Resources: NARE = (AGR + MIN)/YFC, where AGR is GDP originating in agriculture, hunting, forestry, and fishing; MIN is GDP originating in mining and quarrying; and YFC is GDP at factor cost, all in millions of domestic-currency units (OECD [no date] for all OECD countries except the Netherlands; United Nations [1985] for all other countries except Malaysia). For Italy, a joint figure for GDP originating in mining (and quarrying) and manufacturing is allocated in proportion to 1973 values, found in United Nations (1980); for Luxembourg, it is allocated in proportion to 1975 values, found in United Nations (1982). For Malaysia, the UN figures are available for 1973 but not 1975. Estimates for 1975 are obtained from data for that year in World Bank (1983), made consistent with UN data by multiplication by the ratio of the UN to World Bank figure for 1973, for each series.

Literacy Rate: LIT (World Bank 1983). For countries for which a figure for 1975 does not exist, an estimate is obtained by linear interpolation of figures for the years closest to 1975 on either side, except for Austria for which the 1974 figure is taken.

Notes

1. As Clague (1988a, p. 238) writes: “It is obvious, moreover, that the tradability of goods is a matter of degree rather than a simple dichotomy, but theorists need to make simplifications, and the assumption of a tradable/nontradable dichotomy has proven quite useful in thinking about national price levels.”
2. Clague (1985, p. 998) notes: “It is typically assumed that the law of one price holds for tradables but not for nontradables.” Other statements to this effect are made by Clague (1986a, p. 321; 1988a, pp. 237–238); Kravis and Lipsey (1983, p. 11); and Kravis et al. (1978, pp. 218–219).
3. The third proposition is noted by Clague: “The key to the change in *RPL* [the national price level] is found by looking at the relative prices of services and commodities” (1986a, p. 321) and “the relationship of the prices of nontradables to the prices of tradables is at the heart of most available theories of national price levels” (1988a, p. 238). Other statements of this methodology are made by Kravis and Lipsey (1983, pp. 11, 17) and Kravis (1984, p. 29).
4. When a regression is run on *P*, which is rare, it is done only as subsidiary to the principal estimation, as in Kravis and Lipsey (1983, pp. 24–25). Sometimes regressions are run *separately* on the components of *P: PN* (the nontradable price-level ratio) and *PT* (the tradable price-level ratio), as is done by Kravis and Lipsey (1983, pp. 23–24; 1987, pp. 112–114) and Clague (1986a, p. 323). Such bifurcation of *P* makes no sense.
5. While the definitions are unusual, they are consistent with conventional formulations, as for example, in Ruggles (1967, pp. 181–184) and Kravis et al. (1978, pp. 73–76; 1982, pp. 74–77, 89–90).
6. See Kravis et al. (1978, p. 74). In other words, one unit of base-country currency has the same purchasing power as one unit of international currency. This is true only at the GDP level; the base-country/international PPP at any disaggregate level (say, tradables or nontradables) is *not* identically equal to unity, as noted in Kravis et al. (1982, p. 7).
7. Although a linear (log-linear) formulation with a Laspeyres (Walsh) PPP index would be correct, it happens that every econometric study to date has erred in specification of the functional form. For example, Kravis and Lipsey (1983, 1987) and Clague (1986a, 1988a) use linear or log-linear functional forms with the Geary-Khamis index. Salazar-Carrillo (1982b) adopts the Walsh index (see Salazar-Carrillo 1982a), but he uses a linear rather than log-linear functional form!
8. Clague’s (1986c) result holds under specific, but not purely mobile, factors of production. Kravis and Lipsey (1983, p. 16) argue, first, that generally accepted theory offers no support for any unambiguous effect

of a transfer on PL and, second, that true causation is actually the reverse (from PL to the current-account balance.)

9. As Clague (1988a, p. 247) states: "The exchange rate responds more quickly to changes in the money supply than the prices of goods; consequently, a rapid expansion in the money supply should be associated with a lower-than-normal price level."
10. Of course, for the Geary-Khamis index, $(N - 1)$ of each of the variables P_d, SN_d, and Υ_d, indexed by d, would be involved.
11. These influences are in accordance with the literature on the law of one price, as summarized in Officer (1986, pp. 161–163).
12. See Clague (1985, p. 1002; 1986a, p. 320; 1986c, p. 160; 1988a, p. 238).
13. Following Clague (1986a, b), the ICP 34-country sample is reduced to 31 by excluding observations on three centrally-planned economies (Hungary, Poland, and Romania) because of unavailability of data on explanatory variables.
14. The definition of U^2 is $\Sigma(E - A)^2/\Sigma A^2$, where A is the actual and E the estimated value of the variable under consideration. Theil (1971, pp. 26–52) is concerned with time-series forecasting errors. Therefore he defines A and E as percentage changes from the previous period's actual value. With cross-sectional data, this is not possible; so A and E are taken as level, nontransformed variables. A zero value for U corresponds to a perfect fit between E and A.
15. Clague (1985, p. 1005; 1986a; 1988a, pp. 240–241) considers only tourism rather than all services and interprets the variable as having an unambiguously demand effect.
16. Clague (1986a, 1988a) uses alternative variables, each of which covers only part of the natural-resources sector: agricultural land per capita (unambiguously a supply variable) and the share of mineral production in GDP.
17. Explicitly, Kravis and Lipsey (1983, p. 29) state, without empirical testing: "Real income per capita is the major source of variation among countries in the price levels of both tradable and nontradable goods and in the total price level." Implicitly, Clague (1988a, p. 239) writes: "Regression equations explaining national price levels practically always include real per capita income as one of the independent variables. Theoretical attention has therefore turned toward variables that would explain price levels when per capita income is held constant."

References

Bhagwati, Jagdish N. "Why Are Services Cheaper in the Poor Countries?" *Economic Journal* 94 (June 1984): 279–286.

Clague, Christopher K. "A Model of Real National Price Levels." *Southern Economic Journal* 51 (April 1985): 998–1017.
_______. "Determinants of the National Price Level: Some Empirical Results." *Review of Economics and Statistics* 68 (May 1986a): 320–323.
_______. "Short-Cut Estimates of Real Income." *Review of Income and Wealth* 32 (September 1986b): 313–331.
_______. "Tariffs, Transfers and the Real Exchange Rate." *Southern Economic Journal* 53 (July 1986c): 155–169.
_______. "Explanations of National Price Levels." In *World Comparison of Incomes, Prices and Product*, edited by Jorge Salazar-Carrillo and D.S. Prasada Rao, 237–262. Amsterdam: North-Holland, 1988a.
_______. "Purchasing-Power Parities and Exchange Rates in Latin America." *Economic Development and Cultural Change* 36 (April 1988b): 529–541.
Dornbusch, Rudiger. "Expectations and Exchange Rate Dynamics." *Journal of Political Economy* 84 (December 1976): 1161–1176.
International Monetary Fund. *Balance of Payments Yearbook*. Vol. 31. Washington, DC: International Monetary Fund, 1980.
_______. *Balance of Payments Statistics. Yearbook, Part 1*. Vol. 34. Washington, DC: International Monetary Fund, 1983.
Isenman, Paul. "Inter-Country Comparison of 'Real' (PPP) Incomes: Revised Estimates and Unresolved Questions." *World Development* 8 (January 1980): 61–72.
Kravis, Irving B. "Comparative Studies of National Incomes and Prices." *Journal of Economic Literature* 22 (March 1984): 1–39.
_______. "The Three Faces of the International Comparison Project." *World Bank Research Observer* 1 (January 1986): 3–26.
Kravis, Irving B., Alan Heston, and Robert Summers. *International Comparisons of Real Product and Purchasing Power.* Baltimore: Johns Hopkins University Press, 1978.
_______. *World Product and Income*. Baltimore: Johns Hopkins University Press, 1982.
Kravis, Irving B., Zoltan Kenessey, Alan Heston, and Robert Summers. A *System of International Comparisons of Gross Product and Purchasing Power.* Baltimore: Johns Hopkins University Press, 1975.
Kravis, Irving B., and Robert E. Lipsey. *Toward an Explanation of National Price Levels.* Princeton Studies in International Finance, No. 52, International Finance Section, Princeton University, Princeton, NJ, 1983.
_______. "The Assessment of National Price Levels." In *Real-Financial Linkages Among Open Economics*, edited by Sven W. Arndt and J. David Richardson, 97–134. Cambridge, MA: MIT Press, 1987.
OECD. *National Accounts 1970–1982.* Volume II. Washington, DC: OECD, n.d.

Officer, Lawrence H. “The Law of One Price Cannot be Rejected: Two Tests Based on the Tradable/Nontradable Price Ratio.” *Journal of Macroeconomics* 8 (Spring 1986): 159–182.

Ruggles, Richard. “Price Indexes and International Price Comparisons.” In *Ten Economic Studies in the Tradition of Irving Fisher*, 171–205. New York: Wiley, 1967.

Salazar-Carrillo, Jorge. “Comparisons of Purchasing Power and Real Product in Latin America.” Paper presented at the International Association for Research in Income and Wealth, Luxembourg, September 1982a.

_______. “The Purchasing Power Estimation of Equilibrium Exchange Rates.” *Economia Internazionale* 35 (February 1982b): 3–13.

Salter, W.E.G. “Internal and External Balance: The Role of Price and Expenditure Effects.” *Economic Record* 35 (August 1959): 226–238.

Theil, Henri. *Applied Economic Forecasting*. Amsterdam: North-Holland, 1971.

United Nations. *Yearbook of National Accounts Statistics, 1979*. Volume I. New York, NY: United Nations, 1980.

_______. *Yearbook of National Accounts Statistics, 1980. Volume I*, Part 2. New York, NY: United Nations, 1982.

_______. *National Accounts Statistics: Main Aggregates and Detailed Tables, 1982*. New York, NY: United Nations, 1985.

_______. *World Comparisons of Purchasing Power and Real Product for 1980*. Part One. New York, NY: United Nations, 1986.

World Bank. *World Tables: The Third Edition*. Volume I. Baltimore, MD: Johns Hopkins University Press, 1983.

CHAPTER 8

Afterword to Part II

8.1 Absolute and Relative PPP (Chapter 5)

The theory in Chapter 5 clearly relates to the relationship between Cassel's absolute and relative PPP presented in Chapter 2. The empirical analysis in Chapter 5 exhibits positive results for relative PPP, with use of legitimate econometrics outside time-series analysis. Two other comments: First, the samples are over 1950–1975, which means mainly the Bretton Woods period. Adjustably fixed exchange rates–not floating rates–were the norm. Floating exchange rates were considered aberrations, at least until 1973. Second, cost-of-living PPP data are from the German Statistical Office, the source for the Houthakker dollar/mark PPP in Chapter 4.

8.2 Law of One Price (Chapter 6)

The data in Chapter 6 encompass primarily 1950–1975 and reaching 1980—periods of the Bretton Woods system and managed floating. This paper differs from all other testing of the "law of one price" [LOP]. Perhaps for that reason, the paper was ignored in the literature, except for a dismissal on the part of Christopher Clague (1989, p. 378), who writes:

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_8

> There seems to me to be abundant evidence that there are large deviations from the law of one price for tradables...The time-series studies of Isard (1977) and Kravis and Lipsey (1978) provide compelling evidence that the prices of domestic and foreign [tradable] goods can diverge substantially, and Officer's (1986) interesting paper does not persuade me otherwise.

One notes that Chapter 6 acknowledges the Isard and Kravis-Lipsey studies as relevant work. Importantly, a later and excellent survey of the LOP and PPP literature reaches a conclusion opposite from that of Clague and close to the finding of Chapter 6:

> While it is fair to say that a universal consensus may not exist yet, the emerging consensus at the present time is converging toward the view that deviations from the LOP are transitory and therefore the LOP holds in the long run among a broad range of tradable goods and currencies.—Ian W. Marsh, Evgenia Passari, and Lucio Sarno. (2012, p. 213)

8.3 National Price Level (Chapter 7)

Chapter 7 was critiqued vigorously by two economists: Christopher Clague (1989), whose reaction was published in the same issue of *Journal of Macroeconomics*, and Arna Desser (1994), whose assessment appeared five years later. Both authors point out that, providing international prices are exogenous for each country Geary-Khamis becomes a tractable price index. As an index-number issue, their statement is irrefutable. As an empirical matter, the extent to which international prices are given even for economically large countries requires testing, which neither I nor these critics perform.

The critics also take issue with my interpretation of the coefficient on the natural-resource variable, and there is merit in their judgment. Further, Clague questions my rejection of monetary or short-run variables in determination of the non-tradable/tradable price level. Part of his position relies on deviations from the "law of one price," which "law" I defend in Chapter 6. The other part involves the desirability "to 'correct' observed national price levels for temporary disturbances in order to expose the structural influences" (Clague, 1989, p. 378). That justification certainly makes sense, pursuant to the control-variables technique in econometrics.[1]

In contrast to my econometric finding that "the accepted view of per capita income as the predominant explanatory variable receives no support," Desser's empirical results "support the standard finding that real income is a major determinant of the real price level." Even though my regressions are outliers, that in no way inpacts on the eight "issues" of Chapter 7.[2]

Clague and Desser argue that the inclusion of the share of nontradables as an explanatory variable should be based on causal hypotheses, not analytical relationships. Here there is perhaps a miscommunication of the theme of Chapter 7. My purpose was "to take seriously" the assumptions of the theory of the national price level and carry through the implications—analytical and econometric—to the end. Thus only structural variables entered the regressions.

Fundamentally, all three of us would agree on the following propositions. (1) There is a precise analytical relationship among the price level, the nontradable/tradable price level, and the share of nontradables. (2) This relationship is specific to the price index adopted. (3) In general-equilibrium analysis, the shares of nontradables and either price level are jointly determined variables.

But I go further: The ultimate logic of the national-price-level literature requires taking the nontradable/tradable price level (rather than the national price level itself) as *the* dependent variable, meaning that the other variables enter purely as determining variables (and that these variables be structural in nature). Clague judges that including the share of nontradables in the equation is a poor second-best to a general-equilibrium model. Desser argues that problems of simultaneity arise. Neither of these positions hold under the extreme logic of the national-price-level literature.

Interestingly, both critics find worthiness in the proposed estimation technique under *Sixth issue*.[3] "Essentially the procedure 'purges' PL of the arithmetic influence of SN_d and provides a cleaner test of the available theories of the national price level, which are really theories of P" (Clague, 1989, pp. 378–379).[4] "Using Officer's analytical PL-P relationship from Equation (10) to compute a vector of Ps for the dependent variable should remove the arithmetic influences of the share of nontradables…the procedure for testing the real price level relationship by regressing the calculated relative price of nontradables on the independent variable, exclusive of the nontradable share, seems an appropriate choice" (Desser, 1994, pp. 323, 325).[5]

The works of Clague, Desser, and myself are far from the last word on understanding the national price level. Interesting later research involves explanatory variables relating to exchange-rate regime, international financial integration, net foreign assets, and economic freedom.[6]

Notes

1. However, it is arguable that exclusion of control variables implies a "strong" test of a given hypothesis.
2. For an extension of the role of per-capita income, see Jeffrey H. Bergstrand (1991).
3. It is also true that both authors discuss problems with the technique.
4. The comma in the Clague quote is not poetic license on my part. Rather, the comma is Clague's own use and indicates his appreciation of an important theme of Chapter 7.
5. The Desser quote is based in part on her empirical results.
6. See Bergstrand (1991), Christian Broda (2006), Mathias Hoffmann and Peter Tillmann (2012), Jaewoo Lee (2007), and Karam Shaar and Mohamed Ariff (2016).

References

Bergstrand, Jeffrey H. (1991). "Structural Determinants of Real Exchange Rates and National Price Levels: Some Empirical Evidence." *American Economic Review* 81 (March), pp. 325–334.

Broda, Christian (2006). "Exchange Rate Regimes and National Price Levels." *Journal of International Economics* 70 (September), pp. 52–81.

Clague, Christopher (1989). "'The National Price Level: Theory and Estimation': A Comment." *Journal of Macroeconomics* 11 (Summer), pp. 375–381.

Desser, Arna (1994). "Real National Price Levels: An Empirical Reassessment." *Journal of Macroeconomics* 16 (Spring), pp. 313–327.

Hoffmann, Mathias, and Peter Tillmann (2012). "International Financial Integration and National Price Levels: The Role of the Exchange Rate Regime." *Journal of International Money and Finance* 31 (October), pp. 1503–1528.

Lee, Jaewoo (2007). "Transfer Effect in National Price Levels." *Review of World Economics* 143 (October), pp. 534–556.

Marsh, Ian W., Evgenia Passari, and Lucio Sarno (2012). "Purchasing Power Parity in Tradable Goods." In Jessica James, Ian W. Marsh, and Lucio Sarno, *Handbook of Exchange Rates*, pp. 189–220. Hoboken, NJ: Wiley.

Shaar, Karam, and Mohamed Ariff (2016). "Re-examination of Price Level Differentials Using Economic Freedom Index." *Journal of International Trade & Economic Development* 25 (August-September), pp. 880–896.

PART III

Standard of Living

CHAPTER 9

Terms of Trade

Originally published in *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Palgrave Macmillan, 2021, pp. 335–347.

9.1 Prelude

9.1.1 New Historical Balance of Payments

The entirety of Officer (2021) constitutes a new historical annual balance of payments for the United States. The beginning year is 1790, the first year of sufficient data; the ending year is 1919, the first year of the official balance of payments. For the terms of trade, the relevant parts of the balance of payments are the goods account and the services account.

Components of the goods account are merchandise, non-monetary silver, ships, and slaves.[1] Components of the services account are transportation (freight [direct trade and carrying trade], port charges, and charter), travel (oceanic and overland), fares (sailing ships and steamships), financial transactions (marine insurance and bankers' commissions), military expenditures, and non-imported slaves.

The balance of goods is "exports *minus* imports" of goods; the balance of services is exports *minus* imports of services; the balance of goods and services is the algebraic sum of balance of goods and balance of services.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_9

Equivalent terms to “exports *minus* imports” are net exports, credits *minus* debits, net credits, receipts *minus* payments, and net receipts.

Merchandise silver is positive from 1853 onward.[2] Discernible imports of enslaved persons end in 1820 (there are no data on slave exports); but purchases of non-imported slaves (a debit) continue intermittently to 1860.[3] There are sales (exports) of ships but no counted purchases (imports). Transit (overland carrying trade) payments begin in 1886 (Officer 2021, p. 236). Military expenditures abroad (there are no receipts) are recorded for the Mexican–American War (1846–1848) and then continuously from the Spanish-American and Philippine-American Wars (1898 onward).

9.1.2 Openness of Economy

Did the U.S. economy become more or less open over time? The usual way to measure openness is to take ratios of trade to GDP. Previous authors confined trade to merchandise. It is more logical to include trade of both goods and services in the numerator. Officer (2021, p. 336) exhibits the ratios of total goods-and-services exports and total goods-and-services imports to GDP. There is co-movement of the two ratios, and both show a trend decline until about 1910; but the subsequent sharp jump in the export ratio, with World War I, is not followed by imports.

The conventional data and narrative are not inconsistent with these findings.[4] However, the ratios here are naturally higher, with trade inclusive of services. This makes the United States much more of an open economy throughout 1790–1913 than, for example, the figures shown in Lipsey (2000, p. 691).

An even higher degree of openness is achieved by taking the ratio of total (exports *plus* imports) goods-and-services trade to GDP, with both numerator and denominator measured in *real* terms.[5] The series, shown in the fifth column of Table 9.1, is broadly similar to the exports and imports GDP ratios; but the level is higher. As a single and comprehensive measure, it summarizes openness better than the other ratios.

No matter the measure, openness drops sharply in 1808, after the Embargo Act, and never again reaches the high level of the pre-Embargo years.

Table 9.1 Terms of trade and related series

Year	*Terms of trade (1790 = 100)*			*Ratio of goods-and-services trade to GDP*		*Percent of services in real trade*	
	Goods	*Services*	*Goods and services*	*1790 = 100*	*Percent*	*Imports*	*Exports*
1790	100	100	100	100	26.41	14.14	5.17
1791	81	116	85	106	28.08	13.47	5.16
1792	79	112	85	107	28.23	13.83	6.08
1793	93	128	103	110	29.02	15.46	7.28
1794	97	107	105	113	29.96	13.24	8.59
1795	133	86	133	132	34.84	10.71	10.92
1796	137	77	134	139	36.83	11.08	12.56
1797	142	75	138	124	32.66	13.12	14.97
1798	171	70	160	119	31.37	13.51	17.27
1799	159	68	152	127	33.64	13.73	16.77
1800	127	75	130	154	40.77	10.77	11.53
1801	134	72	134	164	43.25	9.89	11.08
1802	123	68	121	133	35.24	11.10	10.78
1803	119	76	119	117	30.85	11.59	9.80
1804	117	73	118	137	36.30	11.39	11.14
1805	118	72	119	156	41.19	10.79	11.86
1806	116	73	118	171	45.16	10.49	10.99
1807	113	82	119	155	41.05	10.72	10.36
1808	98	90	109	66	17.30	8.90	8.91
1809	99	85	106	91	24.09	10.89	8.35
1810	108	81	113	95	25.14	11.02	8.89
1811	111	83	114	74	19.50	12.51	8.27
1812	95	104	104	53	13.93	14.48	8.84
1813	73	120	94	25	6.67	37.49	9.65
1814	69	116	83	22	5.90	29.03	9.10
1815	114	72	113	59	15.48	11.98	10.72
1816	159	50	142	76	20.02	9.54	12.86
1817	166	40	143	67	17.70	8.31	12.61
1818	155	52	140	67	17.69	9.32	13.41
1819	135	45	121	64	16.98	8.23	11.39
1820	118	60	113	67	17.65	8.81	9.46
1821	120	52	113	62	16.27	7.67	8.84
1822	123	49	115	70	18.40	6.06	8.99

(continued)

Table 9.1 (continued)

Year	*Terms of trade (1790 = 100)*			*Ratio of goods-and-services trade to GDP*		*Percent of services in real trade*	
	Goods	*Services*	*Goods and services*	*1790 = 100*	*Percent*	*Imports*	*Exports*
1823	116	53	110	73	19.36	6.58	8.42
1824	135	47	125	70	18.54	6.01	9.10
1825	141	42	127	73	19.32	6.02	9.19
1826	117	51	110	72	19.05	6.56	7.71
1827	114	52	108	74	19.51	6.30	7.25
1828	116	48	108	69	18.18	5.63	7.25
1829	125	47	116	65	17.28	5.59	7.59
1830	125	46	117	67	17.79	4.70	7.08
1831	130	49	124	73	19.19	3.97	7.27
1832	140	43	129	69	18.14	3.89	7.39
1833	145	41	132	65	17.20	4.19	7.79
1834	160	42	143	67	17.61	5.07	8.62
1835	181	38	158	70	18.54	4.67	8.87
1836	174	44	155	72	18.93	5.27	9.52
1837	167	51	152	65	17.23	6.41	9.48
1838	168	55	154	63	16.66	6.81	8.69
1839	181	48	163	67	17.59	6.15	9.49
1840	136	60	128	76	20.00	8.86	7.52
1841	148	45	133	71	18.81	6.77	7.34
1842	135	51	124	68	17.96	7.62	6.88
1843	130	56	119	69	18.31	7.33	5.41
1844	131	54	122	72	18.92	6.00	6.15
1845	131	55	123	69	18.13	5.68	6.71
1846	156	78	147	66	17.54	11.75	7.88
1847	164	77	154	71	18.72	12.45	8.08
1848	151	118	150	76	19.97	10.91	7.51
1849	174	40	157	69	18.28	4.25	9.05
1850	198	47	177	71	18.79	5.84	10.79
1851	177	54	161	77	20.21	9.84	11.16
1852	148	46	134	78	20.59	8.50	9.43
1853	156	44	140	79	20.92	8.78	9.85
1854	174	46	155	73	19.22	8.42	10.25
1855	173	44	155	81	21.52	5.80	9.40

(continued)

Table 9.1 (continued)

Year	*Terms of trade (1790 = 100)*			*Ratio of goods-and-services trade to GDP*		*Percent of services in real trade*	
	Goods	*Services*	*Goods and services*	*1790 = 100*	*Percent*	*Imports*	*Exports*
1856	178	39	156	82	21.67	5.50	9.14
1857	178	38	155	68	18.03	6.52	9.14
1858	172	41	152	67	17.77	6.94	8.99
1859	168	42	147	72	19.04	8.31	8.69
1860	164	49	147	66	17.42	7.88	8.75
1861	170	57	154	50	13.17	9.39	9.23
1862	158	54	142	44	11.68	10.30	7.75
1863	138	49	124	46	12.05	12.22	7.71
1864	131	48	119	41	10.88	15.08	7.54
1865	129	70	122	49	12.82	17.02	5.71
1866	146	63	136	57	15.11	15.11	7.43
1867	171	58	154	52	13.79	14.06	8.17
1868	172	60	155	51	13.49	13.12	8.40
1869	166	57	150	55	14.50	12.74	8.04
1870	159	52	143	60	15.91	12.67	7.50
1871	163	53	146	65	17.10	12.16	7.22
1872	166	53	149	67	17.74	11.63	6.15
1873	170	52	151	66	17.33	11.21	5.40
1874	180	48	158	61	16.10	10.64	5.19
1875	188	48	165	59	15.66	10.31	5.42
1876	177	48	155	60	15.83	10.30	5.16
1877	161	49	143	61	16.05	9.90	4.36
1878	163	47	143	62	16.31	10.80	4.46
1879	172	47	152	59	15.56	11.48	5.06
1880	171	52	153	56	14.83	11.78	5.05
1881	184	52	164	50	13.15	11.41	5.26
1882	189	54	169	47	12.35	11.85	5.72
1883	191	54	169	48	12.69	12.46	6.35
1884	196	54	174	48	12.73	13.04	6.31
1885	199	52	175	50	13.19	12.42	5.73
1886	188	50	165	49	12.99	12.37	5.31
1887	181	50	159	47	12.39	13.07	5.21
1888	194	49	170	45	11.82	11.73	5.28

(continued)

Table 9.1 (continued)

Year	*Terms of trade (1790 = 100)*			*Ratio of goods-and-services trade to GDP*		*Percent of services in real trade*	
	Goods	*Services*	*Goods and services*	*1790 = 100*	*Percent*	*Imports*	*Exports*
1889	176	48	156	47	12.38	11.94	5.31
1890	175	49	155	46	12.09	12.50	4.92
1891	183	53	163	47	12.38	12.88	4.67
1892	178	54	159	47	12.41	12.28	5.15
1893	169	60	153	45	11.99	10.45	6.25
1894	164	48	145	50	13.30	11.13	4.26
1895	175	50	154	47	12.50	11.55	4.43
1896	170	50	151	53	13.92	11.64	4.46
1897	175	48	154	55	14.54	7.87	3.55
1898	173	71	159	53	13.90	12.65	3.51
1899	170	53	152	50	13.27	9.46	3.25
1900	179	61	162	49	13.02	11.22	3.99
1901	184	57	165	48	12.72	10.10	3.60
1902	193	57	171	47	12.31	10.29	3.88
1903	197	52	174	45	11.95	13.40	4.61
1904	194	52	171	48	12.77	14.86	4.29
1905	177	52	157	48	12.77	15.72	4.89
1906	182	54	162	49	13.05	15.96	6.12
1907	183	58	165	47	12.45	17.46	7.58
1908	196	55	172	53	14.01	14.86	5.30
1909	204	58	180	51	13.54	15.67	8.64
1910	206	57	181	52	13.67	16.99	8.41
1911	186	54	165	56	14.76	16.68	6.65
1912	181	54	160	57	14.99	17.38	7.33
1913	191	55	169	56	14.78	17.41	7.98
1914	199	60	177	56	14.83	14.27	6.91
1915	206	90	186	69	18.34	12.67	4.96
1916	215	78	191	70	18.58	13.04	6.17
1917	232	70	196	69	18.30	19.30	8.81
1918	242	107	229	64	16.80	37.17	11.11
1919	227	64	201	71	18.63	26.66	11.38

9.1.3 Goods-Versus-Services Economy

Did the United States transition from a goods economy to a services economy or vice-versa, *relative to the rest of the world* (the U.S. international-transactions partners)? Comparing the balance of services with the balance of goods is one approach. The balance of goods is in Officer (2021, column (2) of Table 5 in chapter 18); and the balance of services is column (3) *minus* column (2). The results are remarkable. Until 1852, the balance of services is in surplus in all but one year (1814), and that year it is trivially negative. From 1862 onward, the balance of services is in continuous deficit. In contrast, the balance of goods is in deficit in all but 11 years 1790–1839. After a run of surpluses in the 1840s, it is in deficit again until 1873—but thereupon, except for 1887–1888, uniformly in surplus. Most tellingly, the balance of goods exceeds the balance of services in only ten years until 1872, but in every year thereafter. The suggestion is that the United States may have had a comparative advantage in services relative to goods until the early 1870s, but that this relationship reversed itself in the final quarter of the nineteenth century.

The switch of the United States from comparative advantage in services to that in goods is also illustrated via comparison of the share of services in real exports of goods and services versus the share of services in real imports of goods and services. These shares presumably shed light on the U.S. goods-services production structure. The underlying real-trade series are constructed in Sect. 9.2, and the shares are exhibited in the sixth and seventh columns of Table 9.1.

The peaks in the import series are associated with the War of 1812 and the end of World War I. Interesting is the result for the antebellum period versus thereafter. During the antebellum period, the share of services in imports exceeds that of exports in only 22 of the 71 years. From 1861 onward, the services proportion of imports is *uniformly* higher than that of exports. The structure of the U.S. economy became strongly goods-oriented rather than services-oriented *relative to its trading partners*, and remained there throughout the postbellum period. The conventional statement that the United States became an “industrial power” (Lipsey, 2000, p. 717) is reflected in the behavior of international transactions.

9.2 Construction of Series

9.2.1 Method

Terms of trade (TT) are defined as the price of exports to the price of imports. TT for goods, services, and combined goods and services (the last by summation) are constructed in five steps.

1. Current-dollar series of six aggregates—exports and imports of goods, services, goods-and-services combined (the last by summation)—are assembled, via summation of individual components.
2. Real (constant [1919]-dollar) series for the same six aggregates are generated, again via summation of individual items in goods and services, now re-expressed in *constant (1919) dollars*. Two alternative techniques are utilized. Either a price index (base 1919 = 1) deflates the current-dollar series, or a direct physical measure of the item is applied. In the latter case, the physical measure is converted to constant-dollar form via linking to the current-dollar figure for 1919. Thus all constant-dollar series are expressed in "millions of 1919 dollars" (corresponding to current-dollar series in "millions of current dollars").
3. Implicit deflators for the six aggregates (goods, services, goods-and-services; exports, imports of each) are calculated as current-dollar/constant-dollar ratios.
4. For the three series–goods, services, goods-and-services combined: TT, base 1919 = 1, are computed as the export/import deflator ratio.
5. Division by the 1790 figure and multiplication by 100 converts the three TT series to base 1790 = 100. Details only of steps (1) and (2), and only for the individual components, warrant discussion.

9.2.2 Goods

Exports of goods are merchandise, nonmonetary silver, and ships; while imports of goods are merchandise, nonmonetary silver, and slaves. Current-dollar series of these components are in Officer (2021, pp. 123–126, 157–158).

Real exports of conventional merchandise are current-dollar exports divided by the price index of merchandise exports (Officer 2021, pp. 141–142). Real exports of merchandise silver are current-dollar exports divided by the price of silver (*Historical Statistics*, 1975, series M270). Real exports of ships are derived physically via tonnage sold to foreigners (Officer 2021, chapter 9, Sect. 1.2). The associated price of ship sales (dollars per ton) is generated in Officer 2021, p. 154.

Real imports of conventional merchandise are current-dollar imports divided by the price index of merchandise imports (Officer 2021, pp. 141–142). Real imports of merchandise silver are current-dollar imports divided by the price of silver. Current-dollar imports of slaves are total revenues from U.S.-ship and foreign-ship imports of slaves, that is, the sum of columns (1)–(3) in Officer 2021, chapter 9, Table 1. I convert year-1807 current-dollar imports to constant (1919) dollars via multiplication by the 1919/1807 ratio of the wholesale price index.[6] For the other years, that resulting figure is multiplied by the ratio of the number of imported slaves to the number in 1807.[7] Thus a physical measure of slave imports extends the "constant-dollar" 1807 figure to the other years; but, as with the current-dollar measure, it does not account for the value of the damaged and destroyed lives of the enslaved persons themselves.

9.2.3 Services

9.2.3.1 Transportation

Current-dollar series of components are in Officer (2021, pp. 229–231, 240). Deflators are developed using four sets of weights: 1860 weights for 1790–1860, 1879 weights for 1860–1879, 1913 weights for 1879–1913, 1919 weights for 1913–1919, with one-year linking of the resulting series segments. The year 1860 is the final antebellum year; 1879 marks the end of the greenback period, and 1913 is the final "normal year" before World War I. The weight of an item is the proportion of its value in total transportation exports or imports, as the case may be, and the corresponding price index is based at 1919 = 1.

For exports, the items and price indexes are (a) <u>merchandise freight earnings:</u> freight rate for merchandise exports; (b) <u>specie freight earnings:</u> specie freight rate; (c) <u>port-costs earnings:</u> wholesale price index; (d) <u>transit earnings:</u> overland-exports freight rate. For imports, they are (a) <u>merchandise freight payments and charter payments:</u> freight rate for merchandise imports; (b) specie freight payments: specie freight rate; (c) port-costs payments: U.K. wholesale price index; (d) transit payments: freight rate for inland imports from Canada.[8]

9.2.3.2 Travel

Current-dollar series of components are in Officer (2021, pp. 249–252). Exports of travel services are deflated by the price index for foreign travel expenditures in the United States (Table 10.1, first column). Imports (U.S. travel expenditures abroad) are deflated by the weighted average of U.S.-resident-travel U.K. and Canada price indexes (Officer 2021, pp. 254–255), with weights the proportions of U.S.-citizen expenditures overseas versus in Canada and Mexico.

9.2.3.3 Passenger-Ship Revenue

Rather than constructing price deflators to be applied to the current-dollar series of passenger-ship revenue (fares—Officer 2021, pp. 284–288), I obtain the corresponding "real" series via physical measure: number of passengers.[9] This yields constant-dollar series of greater precision, albeit the procedure lacks an index-number basis. The number of one-way passengers represents "real" oceanic passenger service (ship revenue). A round-trip counts as two one-way trips.

U.S. steamship constant-dollar revenue from foreign residents is set equal to current-dollar revenue in 1919 and extrapolated to earlier years via number of foreign passengers on U.S. steamships (1919 overlap, of course). Foreign steamship constant-dollar revenue from U.S. residents is similarly derived.

The above technique cannot be applied to sailing-ship revenue, because there is zero such revenue in 1919. The solution is to express sailing-ship revenue in 1860 dollars, then extrapolate to 1919 dollars via an inflator factor. The inflator series is steamship per-passenger revenue, dollars per one-way trip, constructed as the ratio of all steamship revenue (U.S. and foreign steamships, revenue from U.S. and foreign residents) to all steamship passengers (U.S. and foreign steamships, number of U.S. and foreign passengers). The inflator factor is the 1919/1860 ratio of the inflator series.

Thus U.S. sailing-ship constant-dollar revenue from foreign residents is set equal to current-dollar revenue in 1860, extrapolated to other years (forward and back, 1860 overlap) via the number of U.S. sailing-ship foreign passengers (resulting in a constant-1860-dollar series), and multiplied by the inflator factor (for the constant-1919-dollar series). Foreign sailing-ship constant-dollar revenue from U.S. residents is similarly generated.

U.S.-ship constant-dollar revenue from foreign residents is the sum of U.S. steamship and U.S. sailing-ship constant-dollar revenues. Foreign-ship constant-dollar revenue from U.S. residents is the sum of foreign steamship and foreign sailing-ship constant-dollar revenues.

9.2.3.4 Financial Transactions

Current-dollar financial transactions are in Officer (2021, pp. 303–306). Deflators for financial transactions are ideally based on the wage rate in the financial sector of the country providing the services. The second-best is a clerical or white-collar wage of the country, and the third-best the unskilled or manual-worker wage. Where a wage rate is not available (or, for consistency), weekly or annual earnings are adopted.

For exports, the deflator begins with Lebergott's (1996, Table A7, series 46) deflator for personal expenditures on finance 1900–1919. Linked to it via 1900 overlap is a clerical-wage series constructed as follows. Railways clerical average annual earnings 1890–1900 (Douglas, 1930, p. 361) are extended to 1860, 1869, 1879, 1889 (1899 overlap) via weekly wage of clerks and weekly wage in retail trade (Lebergott, 1964, pp. 300–301; Barger, 1960, p. 329).[10] Average hourly earnings of manufacturing production workers (Officer, 2009, p. 166) 1859–1889 interpolates 1861–1868, 1870–1878, 1880–1888. Monthly wage of white-collar workers 1822–1860 (Margo, 2000, Table 5B.4), monthly wage of clerks in iron-producing firms 1800–1822 (Jeffrey F. Zabler, 1972, p. 112), and the David-Solar (1977, p. 59) unskilled wage 1790–1800 extrapolate the series to 1790 (via 1860, 1822, 1800 overlaps).[11]

For imports, a weighted average of men and women clerk's pay for 1911–1919 is generated from data in Guy Routh (1980, pp. 7, 90), using Feinstein's (1995, pp. 264–265) manual-worker earnings as wage interpolator.[12] Linked via 1911 overlap is a clerical-earnings series developed from Jeffrey G. Williamson's (1982, p. 48, series 9H) figures for 15 years over 1781–1911, with intervening years interpolated by Feinstein's (1998, pp. 652–653) manual-worker earnings series.[13] The resulting series for 1790–1919 is denominated in U.S. dollars, via the dollar-pound exchange rate.

9.2.3.5 *Military Expenditures Abroad*

Current-dollar military expenditures abroad are in Officer (2021, pp. 303–306). For the deflator, armed forces overseas are extrapolated from 1917–1919 to 1846–1848, 1898–1916 via 1917 overlap with total military personnel on active duty (*Historical Statistics*, 2006, series Ed26). Constant-dollar is set equal to current-dollar military expenditures abroad in 1919 and extrapolated to the earlier years via armed forces overseas. With no direct data, constant-cost military expenditures are estimated by the number of personnel stationed overseas. The implicit assumption is that the cost of maintaining a soldier overseas in earlier years is the same, adjusted for inflation, as it was in 1919.[14]

9.2.3.6 *Non-Imported Slave Purchases*

Purchases of non-imported slaves, in current dollars, are in Officer 2021, pp. 157–158. Just as for imports, I convert year-1807 current-dollar non-imported slave purchases (in Officer 2021, p. 157) to constant (1919) dollars via multiplication by the 1919/1807 ratio of the wholesale price index. For the other years, the resulting figure is multiplied by the ratio of "the number of U.S-ship purchased slaves not imported into the United States" to "this number in 1807."[15]

9.3 Interpretation

Terms of trade (TT) for goods, services, combined goods and services, are shown in Table 9.1, first, second, third columns. The series are based 1790 = 100.

With TT the ratio of the price of exports to the price of imports; for a given volume of trade, TT increasing enhances the *real income* of the country. Even that hypothesis has been challenged.[16] In placing this chapter in Part III, I am making a stronger statement: TT increasing improves the *standard of living* of the country's population. Reasonable assumptions regarding the ratio of consumption to GDP and the distribution of consumption among the population, both at the margin, are implicitly applied.

Computation of the real-income or standard-of-living effect of a TT change is not attempted here, only a simple statement that a heuristic measure of the strength of the TT change may be given by association with the trade/GDP ratio, presented as an index number and an absolute level in the fourth and fifth columns of Table 9.1. Note that trade/GDP measures *strength* (that is, the *magnitude* of the effect of the TT movement on real income and thereupon on the standard of living), without incorporating *direction*, of the TT movement. Trade/GDP can either amplify or diminish the *implication* of a TT movement, whether the movement is an improvement (increase) or deterioration (decrease) in TT.

Lipsey (2000, p. 718) provides figures for the U.S. TT for 14 periods 1789–1913. His series has three deficiencies, absent here. First, as are all existing historical TT, Lipsey's are confined to merchandise trade.[17] A complete measurement requires that goods involve non-merchandise as well as merchandise items, and that services also are included. Second, Lipsey shows only period averages, whereas a continuous annual series is preferred. Third, as is the norm, Lipsey's TT are essentially a ratio of Laspeyres index numbers with disjoint changes in weights over time. In contrast, I use implicit deflators, which have the advantage of incorporating continuous changing weights of components.

Lipsey (2000, p. 717) writes: "On the whole, the picture is one of long-term improvement in the terms of trade–perhaps an increase of two-thirds from the founding of the country to World War I." The 1904–1913/1789–1798 ratio for his merchandise series is 1.71. For my goods series the ratio is 1.66–almost exactly his "two-thirds" assessment. There is a marked difference for services, with the ratio only 0.57—-a TT *decrease* of over 40%; while goods-and-services together have a ratio of 1.45.

TT of goods have an upward trend from the 1820s—and, after 1866, never fall below 160 (except trivially in 1870). From a global maximum of 128 in 1793, TT of services decrease to 72 in 1815; after which TT are below 60 in all but seven years until World War I. Terms-of-trade movement in conjunction with the balances of goods and services suggests that the United States was fortunate indeed to experience a shift in revealed comparative advantage from services to goods! And U.S. businesspersons and merchants were wise to respond to price incentives in their import–export trade of goods and services.

TT for combined goods-and-services is closer to the goods TT than the services TT. Therefore, in the postbellum period, the effect of TT movements on the standard of living was very probably upward. Furthermore, from 1865, the trade/GDP ratio is at least 12% in all but three years (with two of these observations trivially below, the other only slightly below that level). One can reasonably hypothesize that the standard-of-living effect was not only positive but also "strong."

Notes

1. The preferred term is "enslaved persons," in recognition of the full humanity of the people exploited by the slave trade. I largely retain the old usage, for simplicity and in accordance with the traditional economic-history literature. Merchandise may also be termed "conventional merchandise," for differentiation from nonmonetary silver.
2. Monetary gold and silver do not enter the goods account. For the monetary-versus-nonmonetary distinction of the precious-metal flows, see Officer (2021, ch. 8).
3. See Officer (2021, ch. 9, Sect. 2; ch. 17, Sect. 2).
4. See Lipsey (2000, pp. 685, 690–692) and Edelstein (2006, p. 5–441).
5. Real GDP is the Johnston-Williamson (measuringworth.com) constant-dollar series converted to 1919 dollars. The series of real goods and services (in constant [1919] dollars) are constructed in Sect. 9.2.
6. The extrapolator is inadequate statistically, but little else can be done to denominate slave imports in 1919 dollars—a century after slave imports ended! The base 1807 is logical, because that is the highest-activity slave-import year, whether in numbers or value. More fundamentally, though the technique is a logical way to adjust the data from a balance-of-payments standpoint, what results is an inadequate measure of the cruelty associated with enslavement and sale of human beings. The wholesale price index is from Warren and Pearson (1933, pp. 25–27).
7. The number of imported slaves is the sum of columns (2)–(4) in Officer (2021, chapter 4, Table 1).
8. Freight rates are generated in Officer 2021, pp. 120–121, 136, 226–227. The U.K. wholesale price index has components Board of Trade total index, Sauerbeck-*Statist* overall index, and Gayer-Rostow-Schwartz domestic-and-imported commodities index, linked via overlap years 1846 and 1871. Source is Mitchell (1988, pp. 721, 725, 728–729). The index is converted to dollar denomination via multiplication by the dollar-pound exchange rate (Officer 2021, pp. 36–37).
9. Passenger information is in Officer (2021, chapters 3 and 15).
10. For 1869, 1879, 1889, 1899, weekly wage is the product of hourly wage and hours per week.
11. The Margo and Zabler data are controversial (for discussion and references, see Officer (2009, pp. 144–146, 150–151); but, regarding the white-collar wage, their series are the best available for the purpose at hand. For lack of white-collar wage data in the eighteenth century, the David-Solar unskilled wage is used.

12. Routh's wage data are for 1911–1913, 1924. Feinstein's series interpolates 1914–1919. Routh's numbers of male clerks and of female clerks are linearly interpolated 1912–1919 between 1911 and 1921.
13. The Williamson years are 1781, 1797, 1805, 1810, 1815, 1819, 1827, 1835, 1851, 1861, 1871, 1881, 1891, 1901, 1911.
14. See Officer (2021, chapter 17, Sect. 1).
15. The number is: slaves embarked in Africa less those disembarked revenue-earning in the United States (Officer 2021, chapter 4, Table 1, column (1) *minus* column (2)) plus creole slaves landed in Havana (chapter 4, Table 2, column (5) *plus* column (6)).
16. See Derksen (1980) and, in refutation, Greenfield (1984).
17. The merchandise TT series of *Historical Statistics* (2006, series Ee433, Ee436, Ee439, Ee442) have data sources identical to mine, but do not adjust the underlying price series for the gold dollar, fiscal year, or change in base. See Officer (2021, chapter 7, appendix).

References

David, Paul A., and Peter Solar (1977). "A Bicentenary Contribution to the History of the Cost of Living in America." *Research in Economic History* 2, pp. 1–80.

Derksen, J. B. D. (1980). "Measuring International and Inter-Sectoral Terms of Trade: Some Methodological Issues." *Review of Income and Wealth* 26 (3), pp. 341–349.

Douglas, Paul H. (1930). *Real Wages in the United States, 1890–1926*. Boston: Houghton Mifflin.

Edelstein, Michael (2006). "International Transactions and Foreign Commerce." In *Historical Statistics*, pp. 5–441 to 444.

Feinstein, Charles H. (1995). "Changes in Nominal Wages, the Cost of Living and Real Wages in the United Kingdom over Two Centuries, 1780–1990." In Peter Scholliers and Vera Zamagni, eds., *Labour's Reward*, pp. 3–36, 258–266. Aldershot: Edward Elgar.

———. (1998). "Pessimism Perpetuated: Real Wages and the Standard of Living in Great Britain during and after the Industrial Revolution." *Journal of Economic History* 58 (September), pp. 625–658.

Greenfield, C. C. (1984). "The Terms of Trade Index and Terms of Trade Effect." *Journal of the Royal Statistical Society* 33 (December), pp. 371–379.

Harley, C. Knick (1988). "Ocean Freight Rates and Productivity, 1740–1913: The Primacy of Mechanical Invention Reaffirmed." *Journal of Economic History* 48 (December), pp. 851–876.

[*Historical Statistics*] (1975). Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition*. Washington: Government Printing Office.

———. (2006). Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, Gavin Wright, eds. *Historical Statistics of the United States, Millennial Edition*. Washington: Government Printing Office.

Lebergott, Stanley (1964). *Manpower in Economic Growth*. New York: McGraw-Hill.

———. (1996). *Consumer Expenditures*. Princeton: Princeton University Press.

Lipsey, Robert E. (2000). "U.S. Foreign Trade and the Balance of Payments, 1800–1913." In Stanley L. Engerman and Robert E. Gallman, eds., *The Cambridge Economic History of the United States*, vol. 2, pp. 685–732. Cambridge: Cambridge University Press.

Margo, Robert A. (2000). *Wages and Labor Markets in the United States, 1820–1860*. Chicago: University of Chicago Press.

Mitchell, Brian R. (1988). *British Historical Statistics*. Cambridge: Cambridge University Press.

Officer, Lawrence H. (2009). *Two Centuries of Compensation for U.S. Production Workers in Manufacturing*. New York: Palgrave Macmillan.

———. (2021). *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Palgrave Macmillan.

Routh, Guy (1980). *Occupation and Pay in Great Britain, 1906–79*. London: Macmillan.

Simon, Matthew (1960). "The United States Balance of Payments, 1861–1900." In William N. Parker, ed., *Trends in the American Economy in the Nineteenth Century*, pp. 629–716. Princeton: Princeton University Press.

Warren, George F., and Frank A. Pearson (1933). *Prices*. New York: Wiley.

Williamson, Jeffrey G. (1982). "The Structure of Pay in Britain, 1710–1911." *Research in Economic History* (7), pp. 1–54.

Zabler, Jeffrey F. (1972). "Further Evidence on American Wage Differentials, 1800–1830." *Explorations in Economic History* 10 (fall), pp. 109–117.

CHAPTER 10

Consumer Price Index

10.1 Foreign Travel in United States

Originally published in *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Palgrave Macmillan, 2021, pp. 254–260.

A consumer price index (CPI) for foreign travel in the United States is developed annually for 1790–1919.[1] There are five component indexes, with weights food 40%, alcohol 10%, rent 25%, local transportation 12.5%, intercity transportation 12.5%.[2] The base year is 1919 = 1. For the components, always "preferred data" are retail (consumer, rather than wholesale, price indexes).

10.1.1 Food

The U.S. food price index is a composite series with component indexes linked by year of overlap. Sources are *Historical Statistics* (1975, series E137) 1913–1919, Lebergott (1996, Table A7, series 2) 1900–1913, Douglas (1930, p. 36) 1890–1900, Long (1960, p. 158) 1880–1890, Kloft (1995, p. 278) 1851–1880. For the earlier period, I resort to the food wholesale-price index of Warren and Pearson (1933, pp. 25–26) 1790–1851.[3]

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_10

10.1.2 Alcohol

The retail price index of alcohol for 1900–1919, in Lebergott (1996, Table A7, series 5), begins the series. To extend the series, benchmark figures for 1809, 1834, 1836, 1839, 1844, 1849, 1854, 1859, 1869, 1879, 1889, 1899 are constructed from data in Brady (1966, pp. 106, 108).[4] That exhausts retail information. For inter/extrapolation, I develop a wholesale price index of "spirits" for 1790–1900.[5] This index interpolates the intervening years of the benchmark figures and extrapolates the retail price index back to 1899. Using that year of overlap, the benchmark series extrapolates the retail price index to 1809. The wholesale price index extends the retail price index to 1790 via the 1809 overlap.

10.1.3 Rent

The rent index is a composite of indexes in *Historical Statistics* (1975, series E150) 1913–1919, Rees (1961, p. 74) 1890–1913, Long (1960, p. 159) 1880–1890, Kloft (1995, p. 286) 1851–1880, and the New York City index of Margo (1996, p. 623) 1831–1851, the indexes linked by applicable year of overlap.[6] The construction-cost index (with substitution, combined index) of Adams (1975, pp. 811–812, column (7)) 1790–1831 extrapolates the rent index to 1790 via the 1831 overlap.[7]

10.1.4 Transportation

There are two equally weighted transportation price indexes: purchased local transportation and purchased intercity transportation, obtained for 1900–1919 in Lebergott (1996, Table A7, series 57 and 58).[8] For the earlier years, I generate price series corresponding to the domestic-travel passenger series in Officer (2021, chapter 5): omnibus and street railway, for local transportation; stagecoach, and railroad, for intercity transportation.

The omnibus, which presaged the motorbus, was powered by horses and was an extended stagecoach with fixed routes. It originated in 1811 and ended in 1890; thus an omnibus number-of-passengers series is generated annually for 1811–1889 (Officer 2021, pp. 97–99). For the U.S. omnibus fare, I take the New York fare (dollars per passenger), assembled as 0.125 for 1811–1837, 0.0625 for 1847–1853, 0.06 for

1854–1860, 0.10 for 1865–1889.[9] Using the overall consumer price index (Officer, 2007, pp. 143–144; also Sect. 12.4) as interpolator, I obtain 1838–1846 and 1861–1864.

The street railway ran on rails, just like intercity railroads, but was powered by horses, later by electricity. The street railway is the reason for the demise of the omnibus: "iron rails allowed horses to pull twice the load at twice the speed of omnibuses" (McShane and Tarr, 1997, p. 111). The first street railway appeared in 1852, and I developed a number-of-passengers series for 1852–1900 (Officer, 2021, pp. 99–101). The street-railway fare was typically five cents throughout the nineteenth century, not only in New York but also for the country in general. The fare (again, dollars per passenger) that I adopt is 0.05 for 1852–1863 and 1871–1900, 0.0522 for 1864, 0.06 for 1865–1869, 0.0575 for 1870.[10] Thus there is the fare series 1852–1900.

I generate a stagecoach passenger-miles (superior to number-of-passengers) series beginning 1790. but only for stagecoach activity in the East, and end the series in 1840, prior to significant stagecoach facilities for Western travel (for which data are lacking).[11] The stagecoach fare series requires creativity. Dunbar (1937, p. 750) observes that the "Boston and Hartford" fare was $10 in 1783–1800, the "Boston to New York" fare $11 in 1832. With road-mile distances about 102 (Boston to Hartford) and 220 (Boston Commons to New York Central Park), the cost-per-mile is approximately 9.8 and five cents, respectively.[12] Taylor (1951, p. 142) has the average rate for stagecoach travel in general at about seven cents per mile in 1815–1819, the Boston-Worcester turnpike fare at five cents per mile in the early 1830s (take as 1832) and 6.5 cents per mile in 1850. The concurrence of the Dunbar and Taylor figures for 1832 suggests that the Boston-Worcester fare can represent all fares in 1850. These numbers enable assembly of an average stagecoach fare (dollars per passenger-mile): 0.10 in 1790–1800, 0.07 in 1815–1819, 0.05 in 1832, 0.065 in 1850. I use the consumer price index to interpolate 1801–1814, 1820–1831, 1833–1840.

Intercity railroad service began in 1830, and I construct a passenger-miles series from that year to 1900 (Officer, 2021, pp. 102–103). In contrast to the stagecoach, there are good data for a railroad fare (again, dollars per passenger-mile). The source for 1891–1901 is *Historical Statistics* (2006, series Df955); for 1880 and 1890, Fishlow (1966, p. 585, column (3)), with Adams (1895, pp. 617–618) interpolating 1881–1889. The preceding figures are converted from fiscal year ending June 30 to calendar year 1880–1900. An index of passenger revenue per mile for Vermont railroads (Adams, 1944, p. 49) is obtained for 1857–1881 and extended to 1830 (the first year of railroad service) via North's (1965, p. 245) intermittent freight-rate figures.[13] I take Fishlow's figures (1966, p. 585, column (3)) for 1839, 1849, 1859, 1870 as benchmarks, and use the extended Adams series to convert them from average fiscal year ending November 15 to calendar year, and then to interpolate and extrapolate to complete the series 1830–1900.

Using the omnibus and street-railway number-of-passenger numbers as weights, a price index of local transportation (omnibus and street-railway fares) is constructed for 1852–1889. For 1811–1851 and 1890–1900, the omnibus and street-railway fare series, respectively, are used alone—permissible as both series are identically denominated (dollars per passenger). The resulting 1811–1900 series extrapolates Lebergott's local-transportation series to 1811, via the 1900 overlap.

A similar construction is used for intercity transportation. Using stagecoach and railroad passenger-miles as weights, an index of stagecoach and railroad fares (dollars per passenger-mile) is constructed for 1830–1840. The stagecoach fare alone and railroad fare alone constitute the series for 1790–1829 and 1841–1900.[14] The 1790–1900 series extrapolates Lebergott's intercity-transportation price series to 1790 via the 1900 overlap.

10.1.5 Total Index

For lack of information, I assume that, for foreign travelers, intracity and intercity transportation are of equal importance. So, for 1811–1919, a combined transportation index may be constructed as an equally weighted average of the local and intercity indexes. For 1790–1810, the transportation index is extrapolated via 1811 overlap with the intercity index.

Then the total index for foreign travel in the United States may be constructed as a weighted average of four components: food 40%, alcohol 10%, rent 25%, transportation 25%. The total index and its five ultimate components are presented in Table 10.1.

Table 10.1 Consumer price index for foreign travel in United States

Year	*Total index*	*Component indexes*				
		Food	*Alcohol*	*Rent*	*Local transportation*	*Intercity transporta-tion*
1790	0.77	0.41	0.08	0.37	NA	2.38
1791	0.77	0.39	0.09	0.38	NA	2.38
1792	0.79	0.44	0.10	0.39	NA	2.38
1793	0.82	0.49	0.11	0.40	NA	2.38
1794	0.83	0.53	0.11	0.41	NA	2.38
1795	0.88	0.64	0.12	0.42	NA	2.38
1796	0.95	0.73	0.15	0.53	NA	2.38
1797	0.91	0.64	0.13	0.54	NA	2.38
1798	0.88	0.57	0.13	0.53	NA	2.38
1799	0.88	0.58	0.12	0.50	NA	2.38
1800	0.90	0.62	0.12	0.53	NA	2.38
1801	0.92	0.69	0.13	0.52	NA	2.34
1802	0.74	0.52	0.12	0.46	NA	1.92
1803	0.77	0.53	0.12	0.50	NA	1.96
1804	0.77	0.56	0.11	0.46	NA	1.98
1805	0.79	0.63	0.12	0.46	NA	1.90
1806	0.77	0.59	0.11	0.46	NA	1.92
1807	0.72	0.56	0.11	0.46	NA	1.75
1808	0.72	0.44	0.11	0.56	NA	1.84
1809	0.73	0.51	0.13	0.57	NA	1.73
1810	0.73	0.54	0.14	0.58	NA	1.67
1811	0.75	0.55	0.15	0.58	1.21	1.71
1812	0.74	0.55	0.17	0.58	1.21	1.66
1813	0.81	0.67	0.18	0.55	1.21	1.90
1814	0.86	0.71	0.23	0.59	1.21	1.99
1815	0.85	0.73	0.20	0.71	1.21	1.67
1816	0.81	0.67	0.16	0.64	1.21	1.67
1817	0.82	0.72	0.15	0.61	1.21	1.67
1818	0.79	0.67	0.14	0.61	1.21	1.67
1819	0.72	0.55	0.12	0.51	1.21	1.67
1820	0.66	0.43	0.11	0.53	1.21	1.54
1821	0.64	0.40	0.10	0.51	1.21	1.50
1822	0.64	0.43	0.10	0.46	1.21	1.56
1823	0.62	0.42	0.10	0.48	1.21	1.40
1824	0.60	0.39	0.09	0.50	1.21	1.29
1825	0.60	0.39	0.10	0.45	1.21	1.33
1826	0.60	0.38	0.10	0.45	1.21	1.34
1827	0.60	0.39	0.10	0.46	1.21	1.36
1828	0.59	0.39	0.09	0.47	1.21	1.29

(continued)

Table 10.1 (continued)

Year	*Total index*	*Component indexes*				
		Food	*Alcohol*	*Rent*	*Local transportation*	*Intercity transportation*
1829	0.59	0.39	0.09	0.46	1.21	1.28
1830	0.58	0.37	0.09	0.44	1.21	1.27
1831	0.58	0.38	0.11	0.46	1.21	1.19
1832	0.58	0.39	0.10	0.47	1.21	1.18
1833	0.59	0.39	0.10	0.48	1.21	1.18
1834	0.59	0.36	0.09	0.51	1.21	1.22
1835	0.62	0.42	0.10	0.54	1.21	1.28
1836	0.67	0.50	0.10	0.54	1.21	1.37
1837	0.67	0.52	0.11	0.51	1.21	1.42
1838	0.65	0.50	0.12	0.48	1.13	1.38
1839	0.65	0.49	0.13	0.51	1.08	1.38
1840	0.58	0.40	0.10	0.53	0.96	1.22
1841	0.50	0.35	0.09	0.50	0.92	0.91
1842	0.44	0.31	0.08	0.41	0.82	0.80
1843	0.41	0.30	0.08	0.39	0.70	0.72
1844	0.41	0.28	0.08	0.45	0.67	0.73
1845	0.44	0.33	0.09	0.48	0.64	0.77
1846	0.43	0.33	0.08	0.49	0.60	0.76
1847	0.46	0.38	0.10	0.47	0.61	0.81
1848	0.46	0.34	0.09	0.58	0.61	0.73
1849	0.43	0.34	0.09	0.49	0.61	0.65
1850	0.42	0.33	0.09	0.52	0.61	0.61
1851	0.43	0.33	0.08	0.57	0.61	0.56
1852	0.42	0.34	0.07	0.60	0.60	0.45
1853	0.42	0.34	0.08	0.60	0.57	0.47
1854	0.45	0.39	0.10	0.61	0.54	0.49
1855	0.46	0.41	0.12	0.61	0.53	0.53
1856	0.48	0.40	0.13	0.68	0.53	0.57
1857	0.48	0.43	0.12	0.66	0.53	0.54
1858	0.46	0.39	0.11	0.65	0.52	0.55
1859	0.47	0.40	0.12	0.65	0.51	0.55
1860	0.45	0.40	0.11	0.59	0.51	0.55
1861	0.45	0.39	0.10	0.58	0.51	0.56
1862	0.48	0.42	0.12	0.65	0.53	0.54
1863	0.55	0.52	0.18	0.70	0.55	0.57
1864	0.65	0.69	0.41	0.77	0.56	0.55
1865	0.70	0.70	0.55	0.83	0.59	0.68
1866	0.70	0.70	0.53	0.86	0.59	0.65
1867	0.68	0.68	0.47	0.83	0.59	0.67

(continued)

Table 10.1 (continued)

Year	*Total index*	*Component indexes*				
		Food	*Alcohol*	*Rent*	*Local transportation*	*Intercity transportation*
1868	0.68	0.68	0.35	0.85	0.59	0.66
1869	0.64	0.64	0.24	0.84	0.59	0.66
1870	0.63	0.60	0.22	0.87	0.57	0.63
1871	0.63	0.56	0.20	0.97	0.49	0.62
1872	0.61	0.56	0.20	0.93	0.49	0.62
1873	0.60	0.55	0.20	0.89	0.49	0.63
1874	0.59	0.54	0.20	0.86	0.49	0.64
1875	0.58	0.52	0.22	0.83	0.49	0.66
1876	0.56	0.50	0.22	0.80	0.49	0.66
1877	0.57	0.50	0.21	0.81	0.49	0.64
1878	0.54	0.45	0.20	0.81	0.49	0.64
1879	0.53	0.44	0.19	0.83	0.49	0.57
1880	0.54	0.45	0.20	0.85	0.49	0.57
1881	0.54	0.45	0.19	0.85	0.49	0.54
1882	0.54	0.45	0.19	0.85	0.49	0.56
1883	0.54	0.44	0.20	0.84	0.49	0.57
1884	0.53	0.42	0.20	0.84	0.49	0.54
1885	0.52	0.41	0.20	0.84	0.49	0.52
1886	0.52	0.41	0.20	0.84	0.49	0.53
1887	0.52	0.41	0.19	0.84	0.49	0.53
1888	0.52	0.41	0.20	0.84	0.49	0.56
1889	0.51	0.39	0.19	0.84	0.49	0.55
1890	0.51	0.40	0.17	0.84	0.49	0.52
1891	0.51	0.40	0.17	0.84	0.49	0.51
1892	0.51	0.39	0.17	0.85	0.49	0.50
1893	0.51	0.40	0.17	0.85	0.49	0.49
1894	0.50	0.38	0.18	0.84	0.49	0.48
1895	0.49	0.37	0.19	0.81	0.49	0.48
1896	0.49	0.36	0.19	0.82	0.49	0.48
1897	0.49	0.38	0.18	0.79	0.49	0.48
1898	0.49	0.38	0.19	0.79	0.49	0.47
1899	0.49	0.38	0.19	0.78	0.49	0.47
1900	0.49	0.39	0.19	0.76	0.49	0.48
1901	0.50	0.41	0.20	0.78	0.47	0.47
1902	0.51	0.44	0.20	0.77	0.47	0.47
1903	0.52	0.43	0.20	0.82	0.47	0.48
1904	0.53	0.44	0.20	0.86	0.49	0.48
1905	0.53	0.43	0.19	0.87	0.49	0.48
1906	0.54	0.45	0.19	0.88	0.50	0.49

(continued)

Table 10.1 (continued)

Year	*Total index*	*Component indexes*				
		Food	*Alcohol*	*Rent*	*Local transportation*	*Intercity transportation*
1907	0.56	0.47	0.19	0.92	0.51	0.51
1908	0.55	0.47	0.19	0.89	0.50	0.50
1909	0.56	0.50	0.19	0.87	0.50	0.50
1910	0.58	0.52	0.19	0.89	0.51	0.51
1911	0.57	0.52	0.19	0.87	0.51	0.52
1912	0.58	0.55	0.19	0.87	0.51	0.52
1913	0.59	0.53	0.19	0.90	0.54	0.53
1914	0.60	0.55	0.19	0.90	0.56	0.55
1915	0.60	0.54	0.20	0.90	0.56	0.56
1916	0.64	0.61	0.21	0.91	0.59	0.59
1917	0.73	0.78	0.27	0.91	0.66	0.66
1918	0.86	0.90	0.47	0.92	0.90	0.89
1919	1.00	1.00	1.00	1.00	1.00	1.00

NA = not available

Notes

1. The index was generated to estimate foreign travel expenditures in the United States, for a new U.S. historical balance of payments (Officer, 2021). The author's CPI for the domestic U.S. population (that is, for the traditional coverage of CPI) is shown in Sect. 12.4.
2. Inspiration for the weighting pattern comes from Lebergott (1996, pp. 90; 267 [note 133]).
3. The missing year 1792 is interpolated via Bezanson et al. [BGH] (1936, p. 394), wholesale price index for total foods.
4. Figures are constructed as equally weighted price indexes of "rum and whiskey," "malt liquors," and "vinous liquors." For 1809, only the first two indexes are available; the two-component index is linked to the three-component index via the 1834 overlap.
5. Sources are: for 1860–1900, Hanes (1998, pp. 205–206); for 1790–1860, linked via 1860 overlap, Warren and Pearson (1933, pp. 25–26), with the BGH (1936, p. 393) "wines" index serving to interpolate the missing-year 1792.
6. The Brady rent index (Gallman and Weiss, 1969, p. 292) is rejected, because it behaves contradictorily to the Rees, Long, and Margo indexes.

7. There are two objections to the use of a construction-cost series to represent the rent level. First, rent is determined by the stock of housing, not (per construction cost) the flow (Margo, 1996, p. 605). In response, David and Solar (1977, pp. 44, 56 [note 1]) provide justification that, for the first half of the nineteenth century, rent was closely related to the capitalized value of housing. Second, construction cost does not allow for change in productivity. There are two counters to the latter criticism: The Adams index picks up productivity improvements associated with shifts in factor proportions, and "the sources of productivity improvement in construction do not appear to have been important before the mid-1830s" (Gallman and Rhode, 2019, p. 81).
8. The modes are street railways for local, steam railroads for intercity (Lebergott, 1996, pp. 88, 266 [note 118]).
9. Data are in Taylor (1966a, p. 47), Middleton (1967, p. 13), Williams (1833, p. 160; 1834, p. 172), Glance at New York (1837, p. 260), Holley (1847, p. 55), Foster (1849, p. 63), Disturnell (1876, p. 172), *New York Times* (June 16, 1853, p. 6; December 27, 1856, p. 4; September 30, 1870, p. 4). Rest-of-country fares do not systematically differ from the New York fare.
10. Data sources are Cheape (1980, p. 25), Taylor (1966b, p. 52), Holt (1972, p. 336), Pushkarev, Zupan, and Cumella (1982, p. 5), Durand (1905, p. 10), McShane and Tarr (1997, p. 111), Easton (1859, pp. 22, 24), *New York Times* (various issues). The 1864 and 1870 fares are monthly averages based on the timing of fare changes in New York (*New York Times*, October 12, 1864, p. 4; October 8, 1870, p. 4).
11. The series is exposited in Officer (2021, pp. 101–102).
12. It is good confirmation that the "Boston to New York" $11 fare in 1832 and "this rate being about five cents a mile" are stated by Crocker (1900, pp. 5–6).
13. North's figures are for 1833, 1848, 1851–1852, 1854, 1856, 1858. The Warren-Pearson (1933, pp. 25–26) wholesale price index interpolates 1834–1847 and extrapolates the 1833 figure to 1830–1832. Years 1849–1850, 1853, 1855 are linearly interpolated.
14. Lack of data prevents extending the stagecoach series past 1840. For the same reason, water transportation taken by foreign travelers within the country is not incorporated.

References

Adams, Donald R. Jr. (1975). "Residential Construction Industry in the Early Nineteenth Century." *Journal of Economic History* 35 (December), pp. 794–816.

Adams, Henry C. (1895). "Statistics of Steam Railroad Transportation." In *Report on Transportation Business in the United States at the Eleventh Census*, pp. 591–620. Washington: Government Printing Office.

Adams, Thurston M. (1944). *Prices Paid by Vermont Farmers for Goods and Services and Received by Them for Farm Products, 1790–1940*. Burlington: Vermont Agricultural Experiment Station.

Bezanson, Anne, Robert D. Gray, and Miriam Hussey [BGH]. (1936). *Wholesale Prices in Philadelphia, 1784–1861.* Philadelphia: University of Pennsylvania Press.

Brady, Dorothy S. (1966). "Price Deflators for Final Product Estimates." In Dorothy S. Brady, ed., *Output, Employment, and Productivity in the United States*, pp. 91–115. New York: Columbia University Press.

Cheape, Charles W. (1980). *Moving the Masses*. Cambridge, MA: Harvard University Press.

Crocker, George G. (1900). *From the Stage Coach to the Railroad Train and the Street Car*. Boston: W. B. Clarke.

David, Paul A., and Peter Solar. (1977). "A Bicentenary Contribution to the History of the Cost of Living in America." *Research in Economic History* 2, pp. 1–80.

Disturnell, John. (1876). *New York As It Was and As It Is*. New York: Van Nostrand.

Douglas, Paul H. (1930). *Real Wages in the United States, 1890–1926*. Boston: Houghton Mifflin.

Dunbar, Seymour. (1937). *A History of Travel in America*. New York: Tudor.

Durand, Edward D. (1905). "Street and Electric Railways." In Bureau of the Census, *Street and Electric Railways 1902*, pp. 3–235. Washington: Government Printing Office.

Easton, Alexander. (1859). *A Practical Treatise on Street or Horse-Power Railways*. Philadelphia: Crissy & Markley.

Fishlow, Albert. (1966). "Productivity and Technological Change in the Railroad Sector, 1840–1910." In Dorothy S. Brady, ed., *Output, Employment, and Productivity in the United States*, pp. 583–646. New York: Columbia University Press.

Foster, George C. (1849). *New York in Slices*. New York: W. F. Burgess.

Gallman, Robert E., and Paul W. Rhode. (2019). *Capital in the Nineteenth Century*. Chicago: University of Chicago Press.

Gallman, Robert E., and Thomas J. Weiss (1969). "The Service Industries in the Nineteenth Century." In Victor R. Fuchs, ed., *Production and Productivity in the Service Industries*, pp. 287–352. New York: Columbia University Press.

Glance at New York. (1837). New York: A. Greene.

Hanes, Christopher. (1998). "Consistent Wholesale Price Series for the United States, 1860–1900." In Trevor, J. O. Dick, ed., *Business Cycles since 1820*, pp. 176–209. Cheltenham: Edward Elgar.

[*Historical Statistics*]. (1975). Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition*. Washington: Government Printing Office.

———. (2006). Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, Gavin Wright, eds. *Historical Statistics of the United States, Millennial Edition*. Washington: Government Printing Office.

Holley, Orville L. (1847). *A Description of the City of New York*. New York: J. Disturnell.

Holt, Glen E. (1972). "The Changing Perception of Urban Pathology." In Kenneth T. Jackson and Stanley K. Schultz, eds., *Cities in American History*, pp. 324–343. New York: Alfred A. Knopf.

Kloft, Leonard J. (1995). *Problems and Issues in Constructing National and Regional Price Indices for the United States: 1851–1880*. Ph.D. diss. University of Iowa.

Lebergott, Stanley. (1996). *Consumer Expenditures*. Princeton: Princeton University Press.

Long, Clarence D. (1960). *Wages and Earnings in the United States, 1860–1890*. Princeton: Princeton University Press.

Margo, Robert A. (1996). "The Rental Price of Housing in New York City, 1830–1860." *Journal of Economic History* 56 (September), pp. 605–625.

McShane, Clay, and Joel A. Tarr. (1997). "The Centrality of the Horse in the Nineteenth-Century American City." In Raymond A. Mohl, ed., *The Making of Urban America*, pp. 105–130. Wilmington: Scholarly Resources.

Middleton, William D. (1967). *The Time of the Trolley*. Milwaukee: Kalmbach.

North, Douglass C. (1965). "The Role of Transportation in the Economic Development of North America." In *Les Grandes Voies Maritimes dans le Monde*, pp. 209–246. Paris: S.E.V.P.E.N.

Officer, Lawrence H. (2007). "An Improved Long-Run Consumer Price Index for the United States." *Historical Methods* 40 (summer), pp. 135–147.

———. (2021). *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Palgrave Macmillan.

Pushkarev, Boris S., Jeffrey M. Zupan, and Robert S. Cumella. (1982). *Urban Rail in America*. Bloomington: Indiana University Press.

Rees, Albert. (1961). *Real Wages in Manufacturing, 1890–1914*. Princeton: Princeton University Press.

Taylor, George Rogers. (1951). *The Transportation Revolution*. New York: Holt, Rinehart, and Winston.

———. (1966a). "The Beginnings of Mass Transportation in Urban America: Part I." *Smithsonian Journal of History* 1 (summer), pp. 35–50.

______. (1966b). "The Beginnings of Mass Transportation in Urban America: Part II." *Smithsonian Journal of History* 1 (fall), pp. 31–54.

Warren, George F., and Frank A. Pearson. (1933). *Prices*. New York: Wiley.
Williams, Edwin. (1833, 1834). *New York As It Is*. New York: J. Disturnell.

10.2 The Cost of Living in America: A Political History of Economic Statistics, 1880–2000

This material was originally published in *Journal of American History*, Vol. 97, No. 1 (June 2010), pp. 272–273, and has been reproduced by permission of Oxford University Press.

The Cost of Living in America: A Political History of Economic Statistics, 1880–2000. By Thomas A. Stapleford. (New York: Cambridge University Press, 2009. xviii, 421 pp. Cloth, ISBN 978-0-521-89501-9. Paper, ISBN 978-0-521-71924-7.)

This is purportedly a history of the cost of living (COL). It is that, but it is also undoubtedly a political history of the COL and a history of the federal government bodies responsible for the COL: the Bureau of Labor Statistics (BLS) and its predecessors. As the chapters progress, it is also apparent that the book is an exercise in political economy. Thomas A. Stapleford knows economics as well as political science, and this book will interest academics and government professionals in both fields. I have no doubt that the author's views on the role of economics in government will be considered provocative, perhaps even wrongheaded, by many economists.

The Cost of Living in America consists primarily of three parts, arranged chronologically: 1880–1930 ("Statistics and Labor Reform: Centralization and Its Discontents"), 1930–1960 ("Rationalizing the Democratic Political Order: Cost-of-Living Statistics in the Heart of the New Deal"), and 1960–2000 ("The CPI and the Federal Government: A 'Welfare' Index for the Welfare State"). There is a technical appendix on cost-of-living indices and an epilogue in which the author makes recommendations regarding BLS objectives and operations.

There is much to be praised in the book. The influence of interest groups on technical analysis and the resulting tensions between these groups and between politics and economics are recurring themes. The author shows that the role of COL indices in industrial relations was a forerunner of the functions of the consumer price index (CPI) in income changes directly and via determination of parameters in formulas (for example, poverty thresholds). Stapleford demonstrates conclusively that the COL index and the CPI were buffeted by outside influences (such as labor) from the beginning. The COL index was never a “pure” statistic in any sense. In his well-documented historical narrative, Stapleford effectively challenges those who look at the CPI as only a statistical economic construct. Using political science, he answers the question that he poses early in the work: “How and why did we come to this strange place, where extraordinary amounts of money change hands based on small movements in a controversial and admittedly ambiguous statistic such as the CPI?” (p. 4).

I have some quibbles with the book. A longer historical perspective, with attention to early labor data such as the McLane Report (1833) and the Young Report (1871), would have enriched the book. Quantitative information, such as a time series of the BLS budget and number of staff members, would have been useful. The scattered information about BLS (and its predecessors) leadership should have been expanded and presented in tabular form. But these omissions do not detract from the virtues of what the author has accomplished.

My main concern with the work is implicit in the three parts and explicit in the epilogue. I agree that “the assumption that statistical calculations are straightforward, apolitical facts—cannot be sustained”; but Stapleford goes too far in asserting that “politically important judgments saturate the process of statistical calculation down to a highly detailed methodological level. The strict separation of the political and the technical that is typically used to justify rationalized governance does not exist” (p. 385). Stapleford advocates, in effect, overt politicization of the BLS, with a permanent politically oriented (not technically oriented) advisory committee. In my humble view, the BLS has been too influenced by political considerations—as the author himself demonstrates throughout the book. The BLS needs more objectivity, based on better economic and statistical science. Government is inherently political but better political decisions are obtainable by more—not less—attention to technical analysis.

CHAPTER 11

Compensation of Manufacturing Workers

11.1 Nominal Compensation, Real Compensation, and Standard of Living

Originally published in *Two Centuries of Compensation for U.S. Production Workers in Manufacturing*, Palgrave Macmillan, 2009, pp. 165–180.

11.1.1 Compensation and Its Components

The main results of Officer (2009) are presented in Table 11.1: time series of average hourly compensation (AHC), average hourly earnings (AHE), and average hourly benefits (AHB)—the two latter series constructed in Chapters 5–6 of Officer (2009), the first series their sum. AHB is assumed zero until 1900, then computed for positive values but rounds up to a level of one-tenth of one cent only in 1912. The three variables are rounded to a tenth of a cent (that is, shown to three decimal places) until AHB reaches one cent, which happens in 1936. From then on, the variables are rounded to the nearest cent.

There is a tremendous increase in AHC over the two centuries—understandable because all three variables are measured in nominal (money) terms, that is, they incorporate inflation. The growth in compensation is so great that it can be graphed meaningfully only in logarithmic (ratio)

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_11

Table 11.1 Average hourly compensation, earnings, and benefits: 1800–2006

Year	*AHC*	*AHE*	*AHB*	*Year*	*AHC*	*AHE*	*AHB*
1800	0.040	0.040	0	**1904**	0.152	0.152	0.000
1801	0.040	0.040	0	**1905**	0.156	0.156	0.000
1802	0.044	0.044	0	**1906**	0.163	0.163	0.000
1803	0.044	0.044	0	**1907**	0.173	0.173	0.000
1804	0.046	0.046	0	**1908**	0.163	0.163	0.000
1805	0.047	0.047	0	**1909**	0.167	0.167	0.000
1806	0.046	0.046	0	**1910**	0.175	0.175	0.000
1807	0.046	0.046	0	**1911**	0.178	0.178	0.000
1808	0.047	0.047	0	**1912**	0.187	0.186	0.001
1809	0.048	0.048	0	**1913**	0.197	0.196	0.001
1810	0.046	0.046	0	**1914**	0.199	0.198	0.001
1811	0.051	0.051	0	**1915**	0.200	0.198	0.002
1812	0.052	0.052	0	**1916**	0.237	0.235	0.002
1813	0.050	0.050	0	**1917**	0.285	0.283	0.002
1814	0.051	0.051	0	**1918**	0.358	0.356	0.002
1815	0.051	0.051	0	**1919**	0.431	0.429	0.002
1816	0.049	0.049	0	**1920**	0.539	0.537	0.003
1817	0.047	0.047	0	**1921**	0.483	0.481	0.003
1818	0.047	0.047	0	**1922**	0.444	0.441	0.003
1819	0.045	0.045	0	**1923**	0.481	0.478	0.003
1820	0.044	0.044	0	**1924**	0.507	0.504	0.003
1821	0.050	0.050	0	**1925**	0.503	0.499	0.004
1822	0.046	0.046	0	**1926**	0.510	0.506	0.004
1823	0.046	0.046	0	**1927**	0.516	0.512	0.004
1824	0.049	0.049	0	**1928**	0.519	0.515	0.004
1825	0.048	0.048	0	**1929**	0.516	0.512	0.004
1826	0.051	0.051	0	**1930**	0.527	0.523	0.004
1827	0.050	0.050	0	**1931**	0.513	0.509	0.004
1828	0.048	0.048	0	**1932**	0.446	0.441	0.005
1829	0.055	0.055	0	**1933**	0.441	0.437	0.004
1830	0.057	0.057	0	**1934**	0.527	0.523	0.004
1831	0.056	0.056	0	**1935**	0.542	0.537	0.005
1832	0.052	0.052	0	**1936**	0.55	0.54	0.01
1833	0.057	0.057	0	**1937**	0.63	0.61	0.03
1834	0.052	0.052	0	**1938**	0.64	0.60	0.04
1835	0.054	0.054	0	**1939**	0.64	0.60	0.04
1836	0.052	0.052	0	**1940**	0.67	0.63	0.04
1837	0.061	0.061	0	**1941**	0.74	0.70	0.04
1838	0.058	0.058	0	**1942**	0.86	0.83	0.04
1839	0.058	0.058	0	**1943**	0.98	0.93	0.04

(continued)

Table 11.1 (continued)

Year	*AHC*	*AHE*	*AHB*	*Year*	*AHC*	*AHE*	*AHB*
1840	0.057	0.057	0	**1944**	1.05	1.00	0.05
1841	0.058	0.058	0	**1945**	1.06	1.01	0.05
1842	0.064	0.064	0	**1946**	1.13	1.08	0.05
1843	0.056	0.056	0	**1947**	1.30	1.24	0.06
1844	0.057	0.057	0	**1948**	1.41	1.35	0.06
1845	0.057	0.057	0	**1949**	1.46	1.39	0.07
1846	0.057	0.057	0	**1950**	1.55	1.46	0.09
1847	0.061	0.061	0	**1951**	1.72	1.61	0.11
1848	0.065	0.065	0	**1952**	1.83	1.71	0.12
1849	0.063	0.063	0	**1953**	1.94	1.81	0.13
1850	0.061	0.061	0	**1954**	1.97	1.83	0.14
1851	0.064	0.064	0	**1955**	2.05	1.90	0.15
1852	0.067	0.067	0	**1956**	2.16	1.99	0.16
1853	0.068	0.068	0	**1957**	2.24	2.06	0.18
1854	0.068	0.068	0	**1958**	2.39	2.19	0.20
1855	0.068	0.068	0	**1959**	2.45	2.24	0.22
1856	0.067	0.067	0	**1960**	2.54	2.30	0.24
1857	0.069	0.069	0	**1961**	2.60	2.35	0.25
1858	0.075	0.075	0	**1962**	2.71	2.44	0.28
1859	0.076	0.076	0	**1963**	2.83	2.53	0.29
1860	0.077	0.077	0	**1964**	2.89	2.61	0.29
1861	0.081	0.081	0	**1965**	3.00	2.69	0.32
1862	0.091	0.091	0	**1966**	3.14	2.78	0.35
1863	0.096	0.096	0	**1967**	3.29	2.92	0.37
1864	0.105	0.105	0	**1968**	3.52	3.11	0.41
1865	0.112	0.112	0	**1969**	3.72	3.27	0.45
1866	0.114	0.114	0	**1970**	3.93	3.43	0.49
1867	0.112	0.112	0	**1971**	4.26	3.69	0.57
1868	0.112	0.112	0	**1972**	4.59	3.95	0.64
1869	0.113	0.113	0	**1973**	4.95	4.21	0.74
1870	0.113	0.113	0	**1974**	5.44	4.59	0.85
1871	0.116	0.116	0	**1975**	6.02	5.04	0.98
1872	0.117	0.117	0	**1976**	6.53	5.43	1.11
1873	0.120	0.120	0	**1977**	7.15	5.89	1.26
1874	0.118	0.118	0	**1978**	7.77	6.37	1.40
1875	0.116	0.116	0	**1979**	8.34	6.81	1.53
1876	0.114	0.114	0	**1980**	9.12	7.41	1.71
1877	0.110	0.110	0	**1981**	10.00	8.09	1.91
1878	0.108	0.108	0	**1982**	10.80	8.70	2.10
1879	0.107	0.107	0	**1983**	11.22	9.00	2.22

(continued)

Table 11.1 (continued)

Year	*AHC*	*AHE*	*AHB*	*Year*	*AHC*	*AHE*	*AHB*
1880	0.111	0.111	0	**1984**	11.78	9.41	2.38
1881	0.110	0.110	0	**1985**	12.50	9.94	2.56
1882	0.113	0.113	0	**1986**	12.90	10.21	2.69
1883	0.114	0.114	0	**1987**	13.05	10.35	2.70
1884	0.116	0.116	0	**1988**	13.58	10.68	2.90
1885	0.116	0.116	0	**1989**	14.00	10.95	3.04
1886	0.119	0.119	0	**1990**	14.41	11.25	3.16
1887	0.126	0.126	0	**1991**	14.93	11.57	3.36
1888	0.128	0.128	0	**1992**	15.63	11.95	3.68
1889	0.133	0.133	0	**1993**	16.12	12.17	3.95
1890	0.133	0.133	0	**1994**	16.56	12.40	4.16
1891	0.133	0.133	0	**1995**	16.66	12.67	3.99
1892	0.132	0.132	0	**1996**	16.84	12.97	3.86
1893	0.135	0.135	0	**1997**	18.12	13.99	4.13
1894	0.126	0.126	0	**1998**	18.18	14.20	3.99
1895	0.126	0.126	0	**1999**	18.75	14.70	4.05
1896	0.128	0.128	0	**2000**	19.36	15.17	4.19
1897	0.127	0.127	0	**2001**	19.36	15.29	4.07
1898	0.128	0.128	0	**2002**	21.02	16.47	4.55
1899	0.131	0.131	0	**2003**	21.54	16.65	4.90
1900	0.137	0.137	0.000	**2004**	23.07	17.26	5.81
1901	0.139	0.139	0.000	**2005**	23.92	17.74	6.19
1902	0.148	0.148	0.000	**2006**	24.37	18.33	6.05
1903	0.154	0.154	0.000				

Note AHE and AHB may not sum exactly to AHC, due to rounding

scale, done in Fig. 11.1. Note that equal distances on the vertical axis represent equal percentage (not equal absolute-dollar) increases in AHC.

The composition of AHC is of great-interest. The ratio of benefits to compensation, taken as a percent, is 100 · (AHB/AHC) and plotted in Fig. 11.2. The proportion mark-up of benefits over compensation (AHB/AHC) is different from, and smaller than, the proportion mark-up of benefits over earnings (AHB/AHE), which is used to derive AHB for 1929–2006. Also, the gross-earnings foundation of AHE and the consequent residual concept of AHB imply a lower benefits/compensation ratio than otherwise (see Officer 2009, chapter 1, Gross Earnings versus Regular Earnings; chapter 4, *Average Hourly Benefits*; and chapter 6 , 1929–2006).

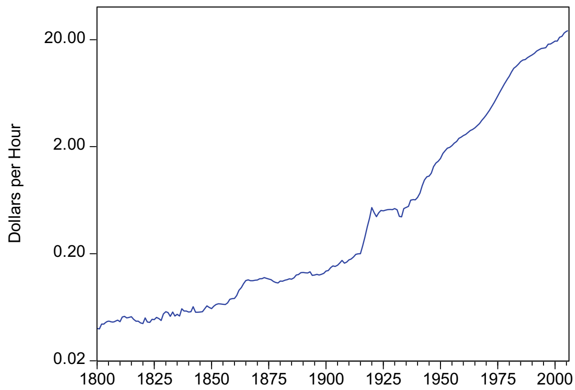

Fig. 11.1 Average hourly compensation (logarithmic scale)

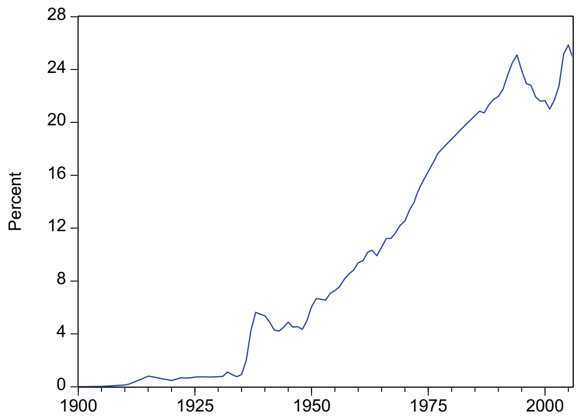

Fig. 11.2 Ratio of benefits to compensation

Until 1900, AHB is so low that it is taken literally as zero. As the graph shows, while the benefits/compensation ratio has an upward trend, the increase is not steady. Benefits reach one percent of compensation only in 1932, fall below that level for three years; exceed five percent in 1938–1940, but fall below five percent in 1941–1949. Benefits first exceed ten percent of compensation in 1962, falling below (but only slightly below) that level only in 1964. In 1984 benefits reach 20% of compensation and never again fall below that figure. The 25-percent level is achieved in 1994 and 2004–2005.

11.1.2 Standard of Living: Alternative Measures

Thus far this chapter has been concerned with the *nominal* wage rate, that is, the wage rate denominated in current dollars. The particular "wage" is AHC, including benefits and expressed in dollars per work-hour. Thus the long-run nominal AHC series is the main contribution of the study.

However, there is an important property of any nominal series: it includes the effect of inflation, and therefore a nominal wage series cannot measure changes in workers' standard of living. From 1800 to 2006 (nominal) average hourly compensation—that is, AHC—increases 608-fold. The corresponding increase in "real" average hourly compensation is far less. The real wage is defined as the nominal wage divided by the consumer price index (CPI). A long-run CP1 series, with reference base 1982–1984 = 100, is developed in Sect. 12.4 and Officer (2008a). Then real average hourly compensation (AHCR) is constructed as AHC/(CPI/100). AHCR is denominated in "1982–1984 dollars per work-hour," listed in Table 11.2, and graphed in Fig. 11.3.

AHCR increases 37-fold from 1800 to 2006, a far lesser magnitude than for nominal compensation. On the one hand, one sees that in earlier years the standard of living of production workers was greater than a comparison of values of the nominal series over time indicates. On the other hand, any CPI series is beset with problems—such as changes in quality of existing commodities, introduction of new commodities, and omission of important commodities—that tend to bias the series upward as one moves forward in time. So there is a sense in which even AHCR understates improvements in the standard of living over time.

Also, it should be remembered that it is the standard of living of *production workers in manufacturing* that is being measured. The CPI series is based on the official Bureau of Labor Statistics (BLS) series for 1917–2006. Until 1978 the official series relates to urban wage-earners and clerical workers. From that date, the series pertains to all urban consumers. To the extent that the consumption pattern of manufacturing production workers differs from the patterns of these groups, the AHCR series incorporates conceptual error. Also, the quality of the CPI series generally deteriorates as one goes backward in time—as it usually does for economic data (including the AHC series).

Table 11.2 Real average hourly compensation: 1800–2006

Years					*Compensation*						
1800–1810	0.33	0.32	0.42	0.40	0.40	0.41	0.39	0.41	0.38	0.40	0.39
1811–1820	0.40	0.40	0.32	0.30	0.34	0.36	0.36	0.38	0.36	0.39	
1821–1830	0.46	0.40	0.45	0.52	0.50	0.53	0.51	0.52	0.61	0.64	
1831–1840	0.67	0.63	0.70	0.63	0.64	0.58	0.66	0.65	0.65	0.68	
1841–1850	0.68	0.81	0.79	0.78	0.78	0.77	0.76	0.85	0.85	0.81	
1851–1860	0.87	0.90	0.91	0.83	0.81	0.81	0.82	0.94	0.94	0.95	
1861–1870	0.95	0.93	0.79	0.69	0.71	0.74	0.78	0.81	0.85	0.90	
1871–1880	0.98	0.99	1.03	1.06	1.09	1.10	1.09	1.11	1.10	1.12	
1881–1890	1.11	1.14	1.18	1.22	1.24	1.30	1.37	1.39	1.49	1.51	
1891–1900	1.50	1.49	1.55	1.51	1.55	1.58	1.57	1.60	1.63	1.69	
1901–1910	1.69	1.78	1.80	1.76	1.83	1.87	1.90	1.83	1.89	1.90	
1911–1920	1.93	1.99	2.05	2.06	2.05	2.23	2.22	2.38	2.49	2.69	
1921–1930	2.70	2.65	2.82	2.97	2.87	2.88	2.97	3.03	3.01	3.16	
1931–1940	3.37	3.27	3.40	3.94	3.95	3.99	4.41	4.54	4.59	4.78	
1941–1950	5.00	5.30	5.64	5.95	5.90	5.79	5.83	5.85	6.13	6.45	
1951–1960	6.64	6.89	7.24	7.33	7.65	7.94	7.95	8.27	8.42	8.56	
1961–1970	8.69	8.96	9.23	9.33	9.52	9.66	9.86	10.11	10.15	10.11	
1971–1980	10.52	10.98	11.14	11.03	11.19	11.48	11.80	11.92	11.50	11.07	
1981–1990	11.00	11.19	11.26	11.34	11.62	11.77	11.49	11.48	11.29	11.03	
1991–2000	10.96	11.14	11.16	11.18	10.93	10.73	11.29	11.16	11.25	11.24	
2001–2006	10.93	11.69	11.71	12.21	12.25	12.09					

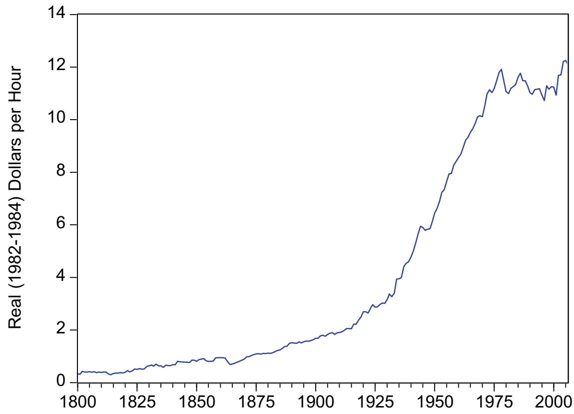

Fig. 11.3 Real average hourly compensation

An alternative measure of the standard of living of manufacturing production workers, original to the present study, is the number of work-hours required to purchase "the consumer bundle." The "value of the consumer bundle" (VCB) is a term invented by the present author (see Sect. 12.3 and Officer 2008b) to describe the "average annual expenditures [per consumer unit]," a BLS series that this author extends back to 1900. VCB emanates from earlier terms—"value of the household bundle" (VHB) and "cost of the (average) household bundle"—developed by Officer and Williamson (2006). VCB is the preferred term, because a "consumer unit" is not the same as a "household." While a household consists of all persons occupying a housing unit, a consumer unit is the decision-making unit for consumer expenditure. Thus a given household can contain more than one consumer unit. This issue, and others relating to the VCB, are discussed in Sect. 12.3 and Officer (2008b).

Let VCB denote the Officer series and HVCB the number of work-hours required to purchase the consumer bundle. For 1900–2006, HVCB is constructed as VCB/AHC. Table 11.3 and Fig. 11.4 ("Required Hours" line) present the HVCB series. Unlike the real wage, standard of living is inversely (rather than directly) related to HVCB. The fewer the number of hours to purchase the consumer bundle, the higher the workers' standard of living. There is a downward trend in HVCB until

Table 11.3 Work-hours required to purchase consumer bundle: 1900–2006

Years	*Number of work-hours*										
1900–1910	5,338	5,599	5,524	5,442	5,618	5,736	5,756	5,619	5,799	6,100	6,020
1911–1920	5,833	5,837	5,659	5,625	5,385	5,248	5,261	4,657	4,358	3,651	
1921–1930	3,286	3,779	3,776	3,551	3,821	3,899	3,796	3,854	3,968	3,466	
1931–1940	3,058	2,814	2,658	2,457	2,555	2,505	2,367	2,272	2,392	2,427	
1941–1950	2,488	2,257	2,127	2,095	2,261	2,698	2,624	2,554	2,429	2,409	
1951–1960	2,284	2,233	2,214	2,230	2,287	2,249	2,267	2,172	2,241	2,221	
1961–1970	2,167	2,138	2,129	2,183	2,205	2,230	2,192	2,182	2,173	2,155	
1971–1980	2,098	2,070	1,921	1,865	1,827	1,825	1,813	1,810	1,842	1,775	
1981–1990	1,699	1,618	1,697	1,865	1,879	1,850	1,871	1,906	1,987	1,969	
1991–2000	1,983	1,909	1,904	1,916	1,936	2,007	1,921	1,954	1,974	1,965	
2001–2006	2,042	1,935	1,895	1,881	1,940	1,986					

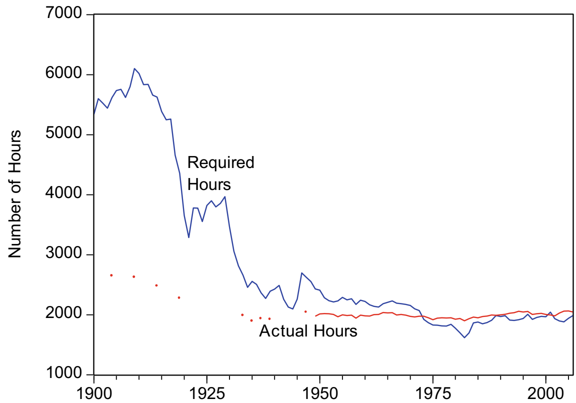

Fig. 11.4 Work-hours: actual and required-to-purchase-consumer-bundle

1982, when the global minimum (1618 hours) occurs, then HVCB increases to 1865 in 1984 and remains within the 1850–2050 range thereafter.

To understand the order of magnitude of the HVCB variable, consider that a 10-hour day (achieved by 1900), 6-day week, and even 52-weeks' work together yield only 3,120 annual work-hours—exceeded by "required work-hours" until 1931. In other words, according to the HVCB measure, the standard of living of the manufacturing production worker was so low in the first three decades of the twentieth century that the fullest-time typical worker could not, by his or her own labor, purchase the consumer bundle! It is also interesting that, while AHCR increases by a multiple of 7.2 over 1900 to 2006, HVCB falls by only a factor of 0.37. For comparison with the AHCR behavior, the inverse of the 0.37 figure is 2.69. Given the criterion of purchasing power over the consumer bundle, AHCR exaggerates the improvement in standard of living by a multiple of more than two-and-a-half.

In the above paragraph, a *hypothetical maximum full-time* work-year provides comparison with the number of work-hours required to purchase the consumer bundle. An alternative comparison measure is the *actual* number of annual work-hours (HACT) per manufacturing production worker. Reliable figures for this variable can be constructed only for certain Census years in the twentieth century: scattered years until 1949 and then continuously. For 1904, 1909, 1914, and 1919, HACT is the product of ADO (average number of days of operation of manufacturing establishments—Officer 2009, Table 5.8) and ADH (average daily hours, using Rees figures—Officer 2009, chapter 3, COMPOSITE SERIES). For 1933, 1935, 1937, 1939, the source of HACT is Census Man-Hour Statistics (sec Officer 2009, chapter 2, *Special Reports* under EARNINGS AND WAGES). For 1933, HACT is the sum of "average hours per wage-earner" in the twelve months; for the other years, HACT is 12 times "average hours per month." For 1947 and 1949–2006, HACT is the ratio of the total hours of production workers (source: Annual Survey of Manufactures—see Officer 2009, chapter 5, 1920–2006) to the average number of production workers (same source).

HACT ("Actual Hours") is plotted along with HVCB in Fig. 11.4. While there is a downward trend in HACT, the trend ends at around 1935—because of missing observations and the limitations of the Man-Hour Statistics themselves, there is an element of uncertainty here—which is much earlier than the corresponding date (1982) for HVCB.

Another innovative standard-of-living measure is the HACT/HVCB ratio: the proportion of the consumer bundle that the typical manufacturing production-worker can purchase from his or her annual earnings. This standard-of-living measure incorporates not only wage but also employment, and is shown in Table 11.4. The actual/required ratio does not exceed fifty percent until 1919, though this milestone could have been reached during the war years (for which data are missing). Not until 1937 is the ratio ever above 80% (with the same caveat of missing observations), and the 90-percent level is reached in 1952–1953 temporarily

Table 11.4 Ratio of actual to consumer-bundle-required work-hours: 1904–2006

Years	*Work-hours actual/required ratio*
1904–1950[a]	0.47 0.43 0.44 0.52 0.75 0.74 0.82 0.80 0.78 0.81 0.84
1951–1960	0.88 0.90 0.91 0.88 0.87 0.88 0.88 0.89 0.89 0.89
1961–1970	0.91 0.94 0.94 0.93 0.92 0.91 0.91 0.92 0.92 0.91
1971–1980	0.94 0.95 1.03 1.05 1.05 1.06 1.08 1.07 1.06 1.08
1981–1990	1.14 1.17 1.14 1.05 1.04 1.06 1.06 1.05 1.00 1.02
1991–2000	1.01 1.06 1.07 1.07 1.06 1.02 1.04 1.03 1.02 1.02
2001–2006	0.98 1.02 1.07 1.09 1.06 1.03

[a]Scattered years, as follows: 1904, 1909, 1914, 1919, 1933, 1935, 1937, 1939, 1947, 1949, 1950

and from 1961 continuously. Only from 1973 onward (with a slight dip in 2001) does the ratio exceed unity. Concretely, only from 1973 does the typical manufacturing production worker have sufficient annual earnings from his or her labor to purchase the entire consumer bundle. Further, in only three years (1981–1983) are annual earnings more than ten percent the cost of the consumer bundle.

It cannot be an exaggeration to state that historically the manufacturing production worker has not been a leading group among consumers in achieving enhancement of standard of living.

11.1.3 Standard of Living: Comparison with Other Studies

Almost every scholar who develops nominal-wage series does so with the ultimate objective of generating corresponding real-wage series or other real-wage information. Therefore application of the real average hourly compensation (AHCR) series of this study to examination of previous historical-studies' conclusions regarding the real wage is instructive. Arbitrarily, a selection is made only from historical studies published after 1965.

Adams (1968, p. 415—see Officer 2009, chapter 2, *Antebellum Records of Firms*) examines changes in real wages in Philadelphia in 1790–1830 to state: "Two periods of rapid increase [in real wage rates] stand out—the 1790's and the period 1815–1830. The real wage increases of the 1790's were largely dissipated by 1815, but from that point on growth was the rule." Adams exhibits the average annual change in real wages of laborers (here representing unskilled occupations) and separately for five skilled occupations over 1790–1815 and 1815–1830. Taking an unweighted average of the results for the skilled occupations and combining the skilled and unskilled figures using the ten-year (1851–1860) Coelho and Shepherd (CS) Northeast weights (see Officer 2009, chapter 5, *Interpolator and Extrapolator Series*), the average annual change in the real wage is 0.39% for 1790–1815 and 4.05% for 1815–1830.

Here the average annual percentage change in any variable Z is computed as $100 \cdot \log(Z_{t+n}/Z_t)/n$, where log represents the natural logarithm, t is the initial year, and $t + n$ the final year. The average annual percentage change in AHCR is 0.31% for 1800–1815 (of necessity, replacing 1790–1815) and 4.15% for 1815–1830—amazingly close to the Adams figures, considering that the Adams Philadelphia data are not utilized in the present study. In all computations in this section (and, in fact, throughout Officer 2009) unrounded figures are used, resulting in superior precision to that provided by rounded figures shown in a table or stated in the text.

Putting to national use his 1821–1860 wage series based on records of civilian Army employees, Margo (2000, p. 109—see Officer 2009, chapter 2, *Records of Civilian Employees of U.S. Army*, and chapter 5, *Interpolator and Extrapolator Series*) estimates the annual growth rate of the U.S. real wage as the coefficient of a time trend, that is, the least-squares estimate of β in the equation $\log W_{RE} = \alpha + \beta \cdot T + \varepsilon$, where W_{RE} is the real wage, T a linear time trend, and ε an error term. Consider Margo's "variable-weights" results (which allow occupation-specific labor-force shares to vary over time in the computation of the real wage—consistent with a current-weight compensation series). Weight Margo's common-laborer and artisan growth rates according to 10-year (1851–1860) Coelho-Shepherd national weights (0.3564, 0.6436)—computed from data in CS 1976, pp. 226, 228). Then the estimated growth rate is 0.84% per year. Applying the same technique and time period to AHCR, the average annual growth rate of the real wage is much greater, at 1.80%.

This divergence in results has several possible interpretations. It is possible that the Margo data underestimate wage growth in the economy at large; it is also possible that the AHCR series overestimates this growth. Perhaps both series are reliable; but, with the Margo series confined to males, the explosive growth in the female wage during this period (see Officer 2009, table 5.12) is incorporated only in AHCR.

Considering the CS real-wage series (see Officer 2009, chapter 5, *Interpolator and Extrapolator Series*, regarding the CS nominal wage), Margo (2000, p. 9) derives an implication for real-wage behavior during the 1850s decade: "the unweighted [Coelho-Shepherd] series suggest that real wages fell during the first half of the 1850s.…Real wages then increased but were no higher in 1860 than in 1851 in any region. Thus, the Weeks Report data suggest that the 1850s was a decade of little or no overall real wage growth."

Although Margo is interpreting certain CS regional series, take here the CS (1976, p. 212) national real-wage series. This series combines all observations, unweighted across occupations and regions. For 1851–1855 the average annual growth rate is −2.29%; for 1856–1860 it is 1.45%. Corresponding figures for AHCR are −1.89 and 3.30%. Thus the

AHCR series is not as pessimistic about the 1850s. In fact, while the CS national real-wage series is 2% lower in 1860 than in 1851, AHCR is 9% higher.

The CS series have an honorable but limited role in developing the AHC (and therefore AHCR) series. The methodological and data differences between AHCR and the CS series are so numerous and substantive that the differences in results are not surprising.

Margo (2006b, p. 2.44) computes a real-wage index for unskilled labor for 1774–1974. He exhibits the series not as a table but only as a graph. Both the numerator (nominal wage) and denominator (CPI) of the Margo series are series constructed by David and Solar (1977, pp. 16–17, 59–60) and reprinted in Margo (2006a) and Lindert and Sutch (2006), respectively. It is interesting that David and Solar themselves do not construct a real-wage series.

Using the time-trend regression technique, Margo estimates the average annual growth rate of that real wage for 1774–1974 (1.5% per year), 1774–1900 (1.2% per year), and 1900–1974 (2.5% per year). Using the same technique, but (of necessity) for 1800–1974, 1800–1900, and 1900–1974, corresponding average annual growth rates for AHCR are 2.0, 1.6, and 2.8% per year. Margo's (2006b, p. 2.44) statement that "two full centuries…over this very long period, real wages have increased substantially" is confirmed—even more so—via the AHCR series. Also substantiated is his observation that "the growth rate of real wages accelerated; growth was slower during the nineteenth century than in the twentieth."

The higher growth rates for AHCR are not surprising, because the David-Solar wage series pertains only to unskilled labor, whereas AHCR incorporates both skilled and unskilled workers. There are other differences between the David-Solar wage series and AHCR, but the directions of their effects are uncertain. Prior to 1890, the David-Solar data are based on unadjusted daily rather than daily-adjusted-to-hourly wage quotations; their series is occupational rather than industry based and so not specific to manufacturing; and, until 1890, their data sources are entirely different from those of AHCR. Inconsequential for the real wage but detracting from direct use is the fact that the David-Solar (nominal) wage series is an index number rather than dollar-denominated. The David-Solar wage series is discussed in David and Solar (1977, pp. 57–68) and Margo (2006a, p. 2.257). There are also conceptual and data differences between the AHCR CPI-component and the David-Solar CPI, discussed in Officer 2008a.

The real-wage growth results of Goldin (2000, p. 565), for 1900–1929 and 1948–1973, are not considered here, because her time dimension of earnings is annual rather than daily or hourly. A comparison with AHCR growth would not be legitimate.

Margo (2006b, p. 2.44) draws the following implication from his graph of the David-Solar real-wage series: "it is apparent that year-to- year (or longer-term) variability in growth rates of real wages—volatility—was very considerable in the nineteenth century but was dampened in the twentieth century." It is not at all clear that this phenomenon is repeated in the AHCR series (Fig. 11.3). In particular, the first half of the twentieth century appears to exhibit cycles not present in the David-Solar series.

To examine relative volatility of the real wage in the two centuries, a technique superior to visual inspection of a graph is to use the Hodrick-Prescott filter to decompose AHCR into trend and cycle. Although Hodrick-Prescott is applied in the same way as in Officer (2009, chapter 5, *Days of Operation*), there are two differences. First, the time period here is 1800–2006. Second, the cyclical component (CAHCR) is defined in the conventional way as AHCR *minus* TAHCR, where TAHCR is the trend component. CAHCR is graphed in Fig. 11.5.

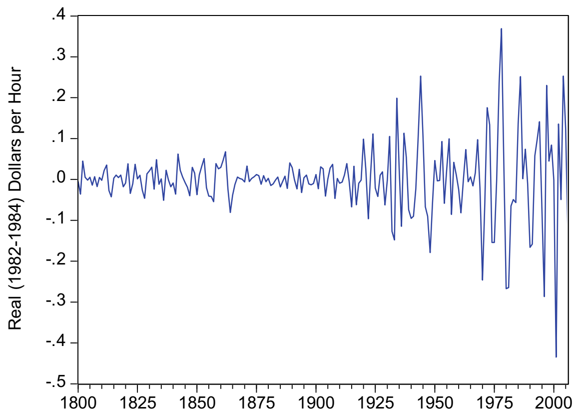

Fig. 11.5 Real average hourly compensation: cycle component

Figure 11.5 shows unambiguously that the cyclical volatility of the real wage AHCR is greater in the twentieth than the nineteenth century—the opposite of Margo's conclusion. Of course, the divergent results are due both to the differing techniques and the different real-wage series.

The relative standard deviations of CAHCR confirm the pattern in Fig. 11.5. For 1800–1899, the standard deviation is 0.027; for 1900–1999, it is 0.111—higher by a factor of 4.1. (The coefficient of variation [ratio of standard-deviation to mean] is not meaningful, because—inherent in the Hodrick-Prescott technique—the mean of CAHCR is zero for the entire time period [1800–2006], and therefore the mean is close to zero for the subperiods.)

In contrast, another of Margo's (2006b, p. 2.44) statements is confirmed using the AHCR series: "the so-called productivity slowdown…began about 1973. A consequence of the slowdown in productivity growth was a marked slowdown in the rate of growth of real wages." Similarly, Goldin (2000, p. 549) notes "labor productivity and real wages lagging in the United States since the mid-1970s." As evidence, Margo examines (separately) the median annual real earnings of male and female full-time workers in the entire economy for 1973–1997. Here, applying the time-trend regression technique to AHCR for 1973–1997, the estimated average annual rate of growth of the real wage is −0.12%. Retardation of real-wage growth during this time period applies, on average, also to manufacturing production workers (males and females together).

11.1.4 Concluding Comments

In summary, and notwithstanding the productivity-slowdown effect on the standard of living, two interesting results follow from historical analysis of the standard of living of the U.S. production worker in manufacturing:

1. Applying the new series of average hourly compensation to the conventional definition of the standard of living—the real wage—the workers' standard of living exhibits greater increases than previous authors have calculated.
2. Applying the new series of average hourly compensation to original and unconventional measures of the standard of living, the increase in workers' standard of living is less impressive—much less impressive—than indicated by the real wage.

References

Adams, Donald R., Jr. (1968). "Wage Rates in the Early National Period: Philadelphia, 1785–1830." *Journal of Economic History* 28 (September), pp. 404–426.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, eds. (2006). *Historical Statistics of the United States, Earliest Times to the Present, Millennial Edition*, 5 vols. New York: Cambridge University Press.

Coelho, Philip R. P., and James F. Shepherd [CS] (1976). "Regional Differences in Real Wages: The United States, 1851–1880." *Journal of Economic History* 13 (April), pp. 203–230.

David, Paul A., and Peter Solar (1977). "A Bicentenary Contribution to the History of the Cost of Living in America." In Paul Uselding, ed., *Research in Economic History* 2, pp. 1–80. Greenwich, CT: JAI Press.

Goldin, Claudia (2000). "Labor Markets in the Twentieth Century." In Stanley L. Engerman and Robert E. Gallman, eds., *The Cambridge Economic History of the United States*, vol. 3, *The Twentieth Century*, pp. 549–623. New York: Cambridge University Press.

Lindert, Peter H., and Richard Sutch (2006a). "Consumer Price Indexes, for All Items: 1774–2003." In *Carter and others*, pp. 3.158–159.

Margo, Robert A. (2000). *Wages and Labor Markets in the United States, 1820–1860*. Chicago: University of Chicago Press.

______ (2006a). "Index of Money Wages for Unskilled Labor: 1774–1974." In Carter and others, pp. 2.256–257.

______ (2006b). "Wages and Wage Inequality." In Carter and others, pp. 2.40–46.

Officer, Lawrence H. (2008a). *The Annual Consumer Price Index for the United States, 1774–2007*. MeasuringWorth website.

______ (2008b). *What Was the Value of the US Consumer Bundle Then?* MeasuringWorth website.

______ (2009). *Two Centuries of Compensation for U.S. Production Workers in Manufacturing*. New York: Palgrave Macmillan.

Officer, Lawrence H., and Samuel H. Williamson (2006). "Better Measurements of Worth." *Challenge: The Magazine of Economic Affairs* 49 (July/August), pp. 86–110.

CHAPTER 12

Afterword to Part III

12.1 Relationship to MeasuringWorth

MeasuringWorth is a historical website created and headed by Samuel H. Williamson. Sam graciously asked me to be co-founder of the site and to serve as Director of Research, which I did for several years. Under that rubric I developed a deep and sustained interest in long-term economic series. Some of the data series in Part III are available on MeasuringWorth in updated format.

12.2 Terms of Trade (Chapter 9)

In reviewing Officer (2021), Devereux (2022) states:

> Take the external terms of trade. Officer covers commodity and service trade for the entire period, where most work in economic history is for commodity trade. He improves deflators and replaces the fixed weight price indices with a more appropriate deflator. The result is that we now have an external terms of trade series for the U.S. from 1790 to now that is superior to the estimates for other developed economies.

Devereux's most-serious criticism is that "some of the most important series appear only as diagrams—including the external terms of trade and the various price series." In listing the terms of trade and related series, Table 9.1 provides a partial response.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_12

12.3 Value of Consumer Bundle

In Officer (2007b) I develop three related U.S. series: value of the consumer bundle (VCB), number of consumer units (CU), and average size of the consumer unit (SZ) annually for 1900–2004. VCB is average annual expenditures per consumer unit. A consumer unit, the entity that makes expenditures decisions, is different from a household. One household is the entirety of persons who occupy a housing unit. There can be more than one consumer unit in a household, and there can be consumer units in a non-household setting, namely, non-institutional "group quarters." So the number of consumer units exceeds the number of households. Table 12.1 presents the series "value of the consumer bundle" (VCB) and "number of consumer units" (CU).

Size of a consumer unit is the number of persons that constitute the unit. Average size of the consumer unit (SZ) is 3.5 1900–1902, 3.4 1903–1917, 3.3 1918, 3.4 1919–1921, 3.3 1922–1933, 3.2 1934–1938, 3.3 1939–1941, 3.2 1942–1962, 3.1 1963–1966, 3.0 1967–1970, 2.9 1971–1974, 2.8 1975–1978, 2.7 1979–1982, 2.6 1983–1991, 2.5 1992–2004.

VCB is denominated in current dollars. To serve as a measure of standard of living over time, VCB needs to be adjusted, performed in Sect. 11.1.2.

One would think that "consumer unit," which by definition is the decision-making unit for expenditures, would be the preferred entity for economic analysis. However, "household," the body of people who occupy a dwelling unit, remains the primary concept for historical research. Consider the monumental work of Robert J. Gordon (2016, p. 36), who computes "average household consumption" [AHC] as $983 in 1870. That figure is too high relative to $733 for VCB in 1900 (the earliest year of the series). How can that be explained?

Gordon (2016, pp. 36; 673, note 1) estimates current-dollar per-capita GDP in a roundabout way, adopts a consumption/GDP ratio of 0.76, and applies a five-person average household, resulting in the $983 figure. What is VCB for 1870? Consider a four-step process.

First, recompute AHC for 1870, retaining Gordon's methodology but using a direct source for per-capita GDP: Louis Johnston and Samuel H. Williamson (2021). The result is $744.[1] This figure is personal consumption expenditures [PCE] divided by number of households.

Table 12.1 Value of consumer bundle and number of consumer units

Year	*VCB (dollars per consumer unit)*	*CU (thousands)*	*Year*	*VCB (dollars per consumer unit)*	*CU (thousands)*
1900	733	21,214	**1953**	4287	49,715
1901	779	21,703	**1954**	4394	50,176
1902	819	22,212	**1955**	4688	50,969
1903	837	22,744	**1956**	4853	51,878
1904	854	23,299	**1957**	5074	52,532
1905	894	23,859	**1958**	5193	53,217
1906	941	24,460	**1959**	5501	54,061
1907	972	25,076	**1960**	5632	55,306
1908	945	25,648	**1961**	5632	55,306
1909	1018	26,229	**1962**	5800	56,753
1910	1055	26,820	**1963**	6014	57,517
1911	1039	27,257	**1964**	6320	58,655
1912	1090	27,713	**1965**	6623	60,203
1913	1116	28,255	**1966**	6994	61,444
1914	1121	28,765	**1967**	7220	62,553
1915	1076	29,147	**1968**	7675	64,416
1916	1243	29,568	**1969**	8087	66,112
1917	1498	29,955	**1970**	8463	67,603
1918	1667	30,117	**1971**	8939	69,145
1919	1878	30,455	**1972**	9512	71,220
1920	1969	31,063	**1973**	9512	71,220
1921	1588	31,804	**1974**	10,147	72,740
1922	1677	32,451	**1975**	11,006	73,914
1923	1817	33,149	**1976**	11,925	75,566
1924	1802	33,883	**1977**	12,960	76,749
1925	1922	34,570	**1978**	14,072	78,534
1926	1988	35,215	**1979**	15,363	79,737
1927	1959	35,827	**1980**	16,184	83,052
1928	1999	36,396	**1981**	16,988	84,249
1929	2046	36,927	**1982**	17,480	85,742
1930	1828	37,411	**1983**	19,043	87,564
1931	1570	37,703	**1984**	21,975	90,223
1932	1255	37,872	**1985**	23,490	91,564
1933	1172	38,263	**1986**	23,866	94,044
1934	1295	38,814	**1987**	24,414	94,150
1935	1385	39,458	**1988**	25,892	94,862

(continued)

Table 12.1 (continued)

Year	*VCB (dollars per consumer unit)*	*CU (thousands)*	*Year*	*VCB (dollars per consumer unit)*	*CU (thousands)*
1936	1385	39,458	**1989**	27,810	95,818
1937	1498	39,391	**1990**	28,381	96,968
1938	1452	39,256	**1991**	29,614	97,918
1939	1526	39,235	**1992**	29,846	100,019
1940	1626	39,203	**1993**	30,692	100,049
1941	1834	39,287	**1994**	31,731	102,210
1942	1950	39,853	**1995**	32,264	103,123
1943	2074	40,289	**1996**	33,797	104,212
1944	2193	40,615	**1997**	34,819	105,576
1945	2401	41,047	**1998**	35,535	107,182
1946	3051	41,955	**1999**	36,995	108,465
1947	3419	42,729	**2000**	38,045	109,367
1948	3597	44,185	**2001**	39,518	110,339
1949	3553	45,858	**2002**	40,677	112,108
1950	3740	47,247	**2003**	40,817	115,356
1951	3938	48,243	**2004**	43,395	116,282
1952	4084	48,988			

Second, correct AHC so the numerator pertains only to the consumer-unit universe. The technique is to multiply AHC by the share of consumer units (population in households *plus* group-quarters residents) in total resident population (PHGQ/POP, in *Adjustment of PCE for consumer-unit universe*, in Officer 2007b, Sect. 5). The data exist for Census years, including 1870.[2] The multiplicative factor is 0.96, the same as for the year 1900, reducing the figure to $715.

Third, estimate the number of consumer units. The technique "to complete the CU series" in Officer (2007b, Sect. 4), was selected there because the developed synthetic series (SCU) is available annually; but there is a serious question of reliability as one proceeds further into the past. A preferred extrapolator, PHGQ (per note 2), can be employed here, because Census data are all that are required. CU in 1870 is estimated as the 1870/1900 PHGQ ratio *times* CU in 1900, with result 11,166 thousand.

Fourth, adjust the corrected AHC so the denominator is the number of consumer units rather than the number of households. With the number of households in 1870 at 7471.754 thousand (Ruggles 2006, Table Ae-A, 1950–1970 definition), the corrective multiplicative factor is 7471.754/11,166, about two-thirds, whence estimated VCB in 1870 is $478.

Conclusion: The Gordon figure for average household consumption in 1870 is *more than double* the VCB for that year! In general, with the number of consumer units exceeding the number of households, average household consumption is an overestimate of the consumer expenditures of decision-making units.

12.4 Consumer Price Index (Chapter 10)

In Officer (2007a) I generate a new U.S. long-run consumer price index (CPI) that is an improvement over alternatives, the most-important of which is the *Historical Statistics* series, presented in Lindert and Sutch (2006). The new series is better in several respects. First, it utilizes a neglected but impressive series of Paul H. Douglas (1930) for the 1914–1917 period. Second, it links component series for conceptual consistency and superior reliability. Third, it embodies enhanced computational accuracy and avoids rounding error. Various tests in Officer (2007a, pp. 141, 145–146) are indicative of the superiority of the new series over the *Historical Statistics* equivalent.

The new CPI series is shown in Table 12.2. This CPI series pertains to the *domestic U.S. population*; it is distinguished from the CPI series in Sect. 10.1, which applies to *foreign travelers in the United States*.

The new CPI improves the official consumer price index, but only within a narrow statistical framework. There are biases (and other limitations) of the CPI that remain in both the official and improved series. For discussion of the biases, one can consult Brent R. Moulton (1996) and David E. Lebow and Jeremy B. Rudd (2003). For the historical political context of the CPI, one may read Thomas A. Stapleford (2009), the subject of the book review in Sect. 10.2. The VCB and (improved) CPI interact in Chapter 11.

Table 12.2 New CPI series

Year	*CPI*	*Year*	*CPI*
1774	7.82	**1890**	8.82
1775	7.41	**1891**	8.82
1776	8.46	**1892**	8.82
1777	10.31	**1893**	8.72
1778	13.38	**1894**	8.34
1779	11.84	**1895**	8.14
1780	13.29	**1896**	8.14
1781	10.72	**1897**	8.04
1782	11.76	**1898**	8.04
1783	10.31	**1899**	8.04
1784	9.91	**1900**	8.14
1785	9.43	**1901**	8.24
1786	9.19	**1902**	8.34
1787	9.02	**1903**	8.53
1788	8.62	**1904**	8.63
1789	8.54	**1905**	8.53
1790	8.86	**1906**	8.72
1791	9.10	**1907**	9.11
1792	9.27	**1908**	8.92
1793	9.59	**1909**	8.82
1794	10.64	**1910**	9.21
1795	12.17	**1911**	9.21
1796	12.81	**1912**	9.40
1797	12.33	**1913**	9.60
1798	11.92	**1914**	9.69
1799	11.92	**1915**	9.74
1800	12.17	**1916**	10.64
1801	12.33	**1917**	12.82
1802	10.39	**1918**	15.06
1803	10.96	**1919**	17.30
1804	11.44	**1920**	20.04
1805	11.36	**1921**	17.90
1806	11.84	**1922**	16.77
1807	11.20	**1923**	17.07
1808	12.17	**1924**	17.10
1809	11.92	**1925**	17.53
1810	11.92	**1926**	17.70
1811	12.73	**1927**	17.37
1812	12.89	**1928**	17.13
1813	15.47	**1929**	17.13

(continued)

Table 12.2 (continued)

Year	*CPI*	*Year*	*CPI*
1814	17.00	**1930**	16.70
1815	14.91	**1931**	15.23
1816	13.62	**1932**	13.66
1817	12.89	**1933**	12.96
1818	12.33	**1934**	13.39
1819	12.33	**1935**	13.73
1820	11.36	**1936**	13.86
1821	10.96	**1937**	14.36
1822	11.36	**1938**	14.09
1823	10.15	**1939**	13.89
1824	9.35	**1940**	14.03
1825	9.59	**1941**	14.73
1826	9.59	**1942**	16.30
1827	9.67	**1943**	17.30
1828	9.19	**1944**	17.60
1829	9.02	**1945**	18.00
1830	8.94	**1946**	19.54
1831	8.38	**1947**	22.34
1832	8.30	**1948**	24.08
1833	8.14	**1949**	23.85
1834	8.30	**1950**	24.08
1835	8.54	**1951**	25.98
1836	9.02	**1952**	26.55
1837	9.27	**1953**	26.75
1838	9.02	**1954**	26.88
1839	9.02	**1955**	26.78
1840	8.38	**1956**	27.18
1841	8.46	**1957**	28.15
1842	7.90	**1958**	28.92
1843	7.17	**1959**	29.16
1844	7.25	**1960**	29.62
1845	7.33	**1961**	29.92
1846	7.41	**1962**	30.26
1847	7.98	**1963**	30.62
1848	7.65	**1964**	31.03
1849	7.41	**1965**	31.56
1850	7.57	**1966**	32.46
1851	7.41	**1967**	33.40
1852	7.49	**1968**	34.80
1853	7.49	**1969**	36.67

(continued)

Table 12.2 (continued)

Year	*CPI*	*Year*	*CPI*
1854	8.14	**1970**	38.84
1855	8.38	**1971**	40.51
1856	8.22	**1972**	41.85
1857	8.46	**1973**	44.45
1858	7.98	**1974**	49.33
1859	8.06	**1975**	53.84
1860	8.06	**1976**	56.94
1861	8.54	**1977**	60.61
1862	9.75	**1978**	65.22
1863	12.17	**1979**	72.57
1864	15.23	**1980**	82.38
1865	15.79	**1981**	90.93
1866	15.39	**1982**	96.50
1867	14.34	**1983**	99.60
1868	13.78	**1984**	103.90
1869	13.21	**1985**	107.60
1870	12.65	**1986**	109.60
1871	11.84	**1987**	113.60
1872	11.84	**1988**	118.30
1873	11.60	**1989**	124.00
1874	11.04	**1990**	130.70
1875	10.64	**1991**	136.20
1876	10.39	**1992**	140.30
1877	10.15	**1993**	144.50
1878	9.67	**1994**	148.20
1879	9.67	**1995**	152.40
1880	9.91	**1996**	156.90
1881	9.91	**1997**	160.50
1882	9.91	**1998**	163.00
1883	9.71	**1999**	166.60
1884	9.51	**2000**	172.20
1885	9.32	**2001**	177.10
1886	9.12	**2002**	179.90
1887	9.22	**2003**	184.00
1888	9.22	**2004**	188.90
1889	8.92	**2005**	195.30

12.5 Compensation of Manufacturing Workers (Chapter 11)

12.5.1 Reception

I was flattered by the comment of Robert E. Hall in the back cover of *Two Centuries of Compensation for U.S. Production Workers in Manufacturing* (Officer 2009): “Highly valuable to scholars interested in quantitative economic history…An intellectual triumph.” Subsequently, Joshua L. Rosenbloom (2009) begins and ends his review of (Officer 2009) as follows:

> I suspect that few people will be tempted to read this slim volume cover to cover. But many of them will find it an extremely valuable reference to which they will return numerous times...Anyone with an interest in the long-run growth of the U.S. economy, or the development of American labor markets will find this book an important and useful reference.”

12.5.2 Data Series

Rosenbloom (2009) makes the following observations on the book’s concluding chapter (which is Chapter 11 of the present work).

> Readers who are interested primarily in the bottom line will want to skip directly to the concluding chapter of this volume, in which the author presents his estimates of average hourly compensation and its components, average hourly earnings, and average hourly benefits in both nominal and real terms. The story that these series tell is in one sense not that surprising. Since 1800, there have been huge increases in nominal compensation; although some of this increase is due to changes in the cost-of-living, real compensation has nonetheless increased dramatically in the last 200 years. The series reported here indicate that average hourly compensation adjusted for inflation increased from $0.33 in 1800 to $12.09 in 2006 (both measured in 1982-84 prices), a nearly 37-fold increase. Growth was somewhat slower in the nineteenth century, and accelerated after 1900, but the series then leveled off in the 1980s, and remained essentially flat until the early 2000s.
>
> While the broad outlines of Officer’s series are consistent with other sources, the shorter run movements of average hourly compensation differ from those of a number of real wage series available over shorter periods. In particular, it appears that average hourly compensation grew faster

than wage series constructed by other scholars for most of the nineteenth century.

John Pencavel (2011, p. 566) observes that my real hourly compensation series (AHCR) "suggests a rise in real hourly compensation between 1890 and 1914 of 36.4%, a figure between Douglas' and Rees' but closer to Rees." He finds that the lower growth in real hourly compensation compared to Rees results from lower growth in nominal compensation (AHC) rather than higher growth in my CPI.

Gordon (2016, p. 279) uses data of Albert Rees (1961) to state: "By 1914 [from 1870], the average nominal manufacturing wage had increased by 30 percent from seventeen cents per hour to twenty-two cents per hour." Consistent with Rosenbloom's rather than Pencavel's comment, my series shows a growth of 45%.

Notes

1. The product of Johnston-Williamson per-capita GDP ($195.76), the Gordon consumption/GDP ratio (0.76), and Gordon's household size (five).
2. Population in households (PH) in Susan Brower and Steven Ruggles (2006, series Ae85), group-quarters residents (GQ) in Steven Ruggles (2006, p. 1–654, Table Ae-A, 1950–1970 definition), resident population (POP) in Michael R. Haines and Richard Sutch (2006, series Aa9). PHGQ = PH + GQ.

References

Brower, Susan, and Steven Ruggles (2006). "Population in Households, by Household Size: 1790–1990." In Carter, Susan B., and others eds., *Historical Statistics of the United States, Millennial Edition*, p. 1-668. Cambridge: Cambridge University Press.

Devereux, John (2022). Review of Lawrence H. Officer, *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Eh.Net Book Review.

Douglas, Paul H. (1930). *Real Wages in the United States, 1890–1926*. Boston: Houghton.

Gordon, Robert J. (2016). *The Rise and Fall of American Growth*. Princeton, NJ: Princeton University Press.

Haines, Michael R., and Richard Sutch (2006). “National Population and the Demographic Components of Change: 1790–2000 [Annual Estimates].” In Carter, Susan B., and others eds., *Historical Statistics of the United States, Millennial Edition*, pp. 1-30 to 34. Cambridge: Cambridge University Press.
Johnston, Louis, and Samuel H. Williamson (2021). “What Was the U.S. GDP Then?” MeasuringWorth.com.
Lebow, David E., and Jeremy B. Rudd (2003). “Measurement Error in the Consumer Price Index: Where Do We Stand?” *Journal of Economic Literature* 41 (March), pp. 159–201.
Lindert, Peter H., and Richard Sutch (2006). “Consumer Price Indexes, for all Items: 1774–2003.” In Carter, Susan B., and others eds., *Historical Statistics of the United States, Millennial Edition*, pp. 3-158 to 159. Cambridge: Cambridge University Press.
Moulton, Brent R. (1996). “Bias in the Consumer Price Index: What Is the Evidence?” *Journal of Economic Literature* 10 (Fall), pp. 159–177.
Officer, Lawrence H. (2007a). “An Improved Long-Run Consumer Price Index for the United States.” *Historical Methods* 40 (Summer), pp. 135–147.
______ (2007b). “Value of the Consumer Bundle: A Data-Series Set.” *Journal of Economic Studies* 34 (No. 3), pp. 160–178.
______ (2009). *Two Centuries of Compensation for U.S. Production Workers in Manufacturing*. New York: Palgrave Macmillan.
______ (2021). *A New Balance of Payments for the United States, 1790–1919: International Movement of Free and Enslaved People, Funds, Goods and Services*. Palgrave Macmillan.
Pencavel, John (2011). “Real Wage Index Numbers.” *American Economic Review* 101 (May), pp. 565–570.
Rees, Albert (1961). *Real Wages in Manufacturing, 1890–1914*. Princeton: Princeton University Press.
Rosenbloom, Joshua L. (2009). Review of Lawrence H. Officer, *Two Centuries of Compensation for U.S. Production Workers in Manufacturing*. EH.Net.
Ruggles, Steven (2006). “Family and Household Composition.” In Carter, Susan B., and others eds., *Historical Statistics of the United States, Millennial Edition*, pp. 1-653 to 659. Cambridge: Cambridge University Press.
Stapleford, Thomas A. (2009). *The Cost of Living in America: A Political History of Economic Statistics, 1880–2000*. New York: Cambridge University Press.

PART IV

Fixed-Rate Monetary Standards

CHAPTER 13

Metallic Standards

Originally published in *The New Palgrave Dictionary of Economics*, third edition. Palgrave Macmillan, 2018: "Gold Standard," pp. 5352–5361; "Silver Standard." pp. 12321–12324; "Bimetallism," pp. 946–951.

13.1 Gold Standard

The classical gold standard (which ended in 1914) and the interwar gold standard are examined within the same framework, but their experiences are vastly different.

13.1.1 Types of Gold Standard

All gold standards involve (a) a fixed gold content of the domestic monetary unit, and (b) the monetary authority both buying and selling gold at the mint price (the inverse of the gold content of the monetary unit), whereupon the mint price governs in the marketplace. A 'coin' standard has gold coin circulating as money. Privately owned bullion (gold in form other than domestic coin) is convertible into gold coin, at (approximately) the mint price, at the government mint or central bank. Private parties may melt domestic coin into bullion–the effect is as if coin were sold to the monetary authority for bullion. The authority could sell gold bars directly for coin, saving the cost of coining.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_13

Under a pure coin standard, gold is the only money. Under a mixed standard, there are also notes issued by the government, central bank, or commercial banks, and possibly demand deposits. Government or central-bank notes (and central-bank deposit liabilities) are directly convertible into gold coin at the fixed price on demand. Commercial-bank notes and demand deposits are convertible into gold or into gold-convertible government or central-bank currency. Gold coin is always exchangeable for paper currency or deposits at the mint price. Two-way transactions again fix the currency price of gold at the mint price.

The coin standard, naturally 'domestic', becomes 'international' with freedom of international gold flows and of foreign-exchange transactions. Then the fixed mint prices of countries on the gold standard imply a fixed exchange rate (mint parity) between their currencies.

A 'bullion' standard is purely international. Gold coin is not money; the monetary authority buys or sells gold bars for its notes. Similarly, a 'gold-exchange' standard involves the monetary authority buying and selling not gold but rather gold-convertible foreign exchange (the currency of a country on a gold coin or bullion standard).

For countries on an international gold standard, costs of importing and exporting gold give rise to 'gold points', and therefore a 'gold-point spread', around the mint parity. If the exchange rate, number of units of domestic per unit of foreign currency, is greater (less) than the gold export (import) point, arbitrageurs sell (purchase) foreign currency at the exchange rate and also obtain (relinquish) foreign currency by exporting (importing) gold. The domestic-currency cost of the transaction per unit of foreign currency is the gold export (import) point; so the 'gold-point arbitrageurs' receive a profit proportional to the exchange-rate/gold-point divergence. However, the arbitrageurs' supply of (demand for) foreign currency returns the exchange rate to below (above) the gold export (import) point. Therefore perfect arbitrage would keep the exchange rate within the gold-point spread. What induces gold-point arbitrage is the profit motive and *the credibility of the monetary-authorities' commitment* to (a) the fixed gold price and (b) freedom of gold and foreign-exchange transactions.

A country can be effectively on a gold standard even though its legal standard is bimetallism. This happens if the gold-silver mint-price ratio is greater than the world price ratio. In contrast, even though a country is legally on a gold standard, its government and banks could 'suspend specie payments', that is, refuse to convert their notes into gold; so that the country is in fact on a 'paper standard'.

13.1.2 Countries on the Classical Gold Standard

Britain, France, Germany and the United States were the 'core countries' of the gold standard. Britain was the 'center country', indispensable to the spread and functioning of the standard. Legally bimetallic from the mid-thirteenth century, Britain switched to an effective gold standard early in the eighteenth century. The gold standard was formally adopted in 1816, ironically during a paper-standard regime (Bank Restriction Period). The United States was legally bimetallic from 1786 and on an effective gold standard from 1834, with a legal gold standard established in 1873–1874–also during a paper standard (the greenback period). In 1879 the United States went back to gold, and by that year not only the core countries but also some British Dominions and noncore western European countries were on the gold standard. As time went on, a large number of other countries throughout the globe adopted gold; but they (along with the Dominions) were in 'the periphery'–acted on rather than actors–and generally (except for the Dominions) not as committed to the gold standard.

Almost all countries were on a mixed coin standard. Some periphery countries were on a gold-exchange standard, usually because they were colonies or territories of a country on a coin standard.

In 1913, the only countries not on gold were traditional silver-standard countries (Abyssinia, China, French Indochina, Hong Kong, Honduras, Morocco, Persia, Salvador), some Latin American paper-standard countries (Chile, Colombia, Guatemala, Haiti, Paraguay), and Portugal and Italy (which had left gold but 'shadowed' the gold standard, pursuing policies as if they were gold-standard countries, keeping the exchange rate relatively stable).

13.1.3 Elements of Instability in Classical Gold Standard

Three factors made for instability of the classical gold standard. First, the use of foreign exchange as official reserves increased as the gold standard progressed. While by 1913 only Germany among the core countries held any measurable amount of foreign exchange, the percentage for the rest of the world was double that for Germany. If there were a rush to cash in foreign exchange for gold, reduction of the gold of reserve-currency countries would place the gold standard in jeopardy.

Second, Britain was in a particularly sensitive situation. In 1913, almost half of world foreign-exchange reserves was in sterling, but the Bank of England had only three percent of gold reserves. The Bank of England's 'reserve ratio' (ratio of 'official reserves' to 'liabilities to foreign monetary authorities held in London financial institutions') was only 31%, far lower than those of the monetary authorities of the other core countries. An official run on sterling could force Britain off the gold standard. Private foreigners also held considerable liquid assets in London, and could themselves initiate a run on sterling.

Third, the United States was a source of instability to the gold standard. Its Treasury held a high percentage of world gold reserves (in 1913, more than that of the three other core countries combined). With no central bank and a decentralized banking system, financial crises were more frequent and more severe than in the other core countries. Far from the United States assisting Britain, gold often flowed from the Bank of England to the United States, to satisfy increases in US demand for money. In many years the United States was a net importer rather than exporter of capital to the rest of the world—the opposite of the other core countries. The political power of silver interests and recurrent financial panics led to imperfect credibility in the US commitment to the gold standard. Indeed, runs on banks and on the Treasury gold reserve placed the US gold standard near collapse in the 1890s. The credibility of the Treasury's commitment to the gold standard was shaken; twice the US gold standard was saved only by cooperative action of the Treasury and a bankers' syndicate, which stemmed gold exports.

13.1.4 Automatic Force for Stability: Price Specie-Flow Mechanism

The money supply is the product of the money multiplier and the monetary base. The monetary authority alters the monetary base by changing its gold holdings and domestic assets (loans, discounts, and securities). However, the level of its domestic assets is dependent on its gold reserves, because the authority generates demand liabilities (notes and deposits) by increasing its assets, and convertibility of these liabilities must be supported by a gold reserve. Therefore the gold standard provides a constraint on the level (or growth) of the money supply.

Further, balance-of-payments surpluses (deficits) are settled by gold imports (exports) at the gold import (export) point. The change in the money supply is the product of the money multiplier and the gold flow, providing the monetary authority does not change its domestic assets. For a country on a gold-exchange standard, holdings of foreign exchange (a reserve currency) take the place of gold.

A country experiencing a balance-of-payments deficit loses gold and its money supply decreases *automatically*. Money income contracts and the price level falls, thereby increasing exports and decreasing imports. Similarly, a surplus country gains gold, exports decrease, and imports increase. In each case, balance-of-payments equilibrium is restored via the current account, the 'price specie-flow mechanism'. To the extent that wages and prices are inflexible, movements of real income in the same direction as money income occur; the deficit country suffers unemployment, while the payments imbalance is corrected.

The capital account also acts to restore balance, via interest-rate increases in the deficit country inducing a net inflow of capital. The interest-rate increases also reduce real investment and thence real income and imports. The opposite occurs in the surplus country.

13.1.5 Rules of the Game

Central banks were supposed to reinforce (rather than 'sterilize') the effect of gold flows on the monetary base, thereby enhancing the price specie-flow mechanism. A gold outflow decreases the international assets of the central bank and the money supply. The central-bank's 'proper' response is: (1) decrease lending and sell securities, thereby decreasing domestic assets and the monetary base; (2) raise its 'discount rate', which induces commercial banks to adopt a higher reserves-deposit ratio, thereby reducing the money multiplier. On both counts, the money supply is further decreased. Should the central bank increase its domestic assets when it loses gold, it engages in sterilization of the gold flow, violating the 'rules of the game'. The argument also holds for gold inflow, with sterilization involving the central bank decreasing its domestic assets when it gains gold.

Monetarist theory suggests the 'rules' were inconsequential. Under fixed exchange rates, gold flows adjust money supply to money demand; the money supply is not determined by policy. Also, prices, interest rates, and incomes are determined worldwide. Even core countries can influence these variables domestically only to the extent that they help determine them in the global marketplace. Therefore the price specie-flow and like mechanisms cannot occur. Historical data support this conclusion: gold flows were too small to be suggestive of these processes; and, at least among the core countries, prices, incomes, and interest rates moved closely in correspondence, contradicting the specie-flow mechanism and rules of the game.

Rather than rule (1), central-bank domestic and international assets moving in the same direction, the opposite behaviour—sterilization—was dominant, both in core and non-core European countries. The Bank of England followed the rule more than any other central bank, but even so violated it more often than not!

The Bank of England did, in effect, manage its discount rate ('Bank Rate') in accordance with rule (2). The Bank's primary objective was to maintain convertibility of its notes into gold, and its principal tool was Bank Rate. When the Bank's 'liquidity ratio' (ratio of gold reserves to outstanding note liabilities) decreased, it usually increased Bank Rate. The increase in Bank Rate carried with it market short-term interest rates, inducing a short-term capital inflow and thereby moving the exchange rate away from the gold-export point. The converse also held, with a rise in the liquidity ratio generating a Bank Rate decrease. The Bank was constantly monitoring its liquidity ratio, and in response altered Bank Rate almost 200 times over 1880–1913.

While the Reichsbank also generally moved its discount rate inversely to its liquidity ratio, other central banks often violated rule (2). Discount-rate changes were of inappropriate direction, or of insufficient magnitude or frequency. The Bank of France kept its discount rate stable, choosing to have large gold reserves, with payments imbalances accommodated by fluctuations in its gold rather than financed by short-term capital flows. The United States, lacking a central bank, had no discount rate to use as a policy instrument.

13.1.6 Reason for Stability: Credible Commitment to Convertibility

From the late 1870s onward, there was absolute private-sector credibility in the commitment to the fixed domestic-currency price of gold on the part of Britain, France, Germany, and other important European countries. For the United States, this absolute credibility applied from about 1900. That commitment had a contingency aspect: convertibility could be suspended in the event of dire emergency; but, after normal conditions were restored, convertibility and honoring of gold contracts would be re-established at the pre-existing mint price—even if substantial deflation was required to do so. The Bank Restriction and greenback periods were applications of the contingency. From 1879, the 'contingency clause' was exercised by none of these countries.

The absolute credibility in countries' commitment to convertibility at the existing mint price implied that there was zero 'convertibility risk' (Treasury or central-bank notes non-redeemable in gold at the established mint price) and zero 'exchange risk' (alteration of mint parity, institution of exchange control, or prohibition of gold export).

Why was the commitment to credibility so credible?

1. Contracts were expressed in gold; abandonment of convertibility meant violation of contracts—anathema to monetary authorities.
2. Shocks to economies were infrequent and generally mild.
3. The London capital market was the largest, most open, most diversified in the world, and its gold market was also dominant. A high proportion of world trade was financed in sterling, London was the most important reserve-currency center, and payments imbalances were often settled by transferring sterling assets rather than gold. Sterling was an international currency—a boon to other countries, because sterling involved positive interest return, and its transfer costs were much less than those of gold. Advantages to Britain were the charges for services as an international banker, differential interest return on its financial intermediation, and the practice of countries on a sterling (gold-exchange) standard of financing payments surpluses with Britain by piling up short-term sterling assets rather than demanding Bank gold
4. 'Orthodox metallism'–authorities' commitment to an anti-inflation, balanced-budget, stable-money policy—reigned. This ideology implied low government spending, low taxes, and limited monetization of government debt. Therefore, it was not expected that a country's price level would get out of line with that of other countries.
5. Politically, gold had won over paper and silver, and stable-money interests (bankers, manufacturers, merchants, professionals, creditors, urban groups) over inflationary interests (farmers, landowners, miners, debtors, rural groups).

6. There was a competitive environment and freedom from government regulation. Prices and wages were flexible. The core countries had virtually no capital controls, Britain had adopted free trade, and the other core countries had only moderate tariffs. Balance-of-payments financing and adjustment were without serious impediments.
7. With internal balance an unimportant goal of policy, preservation of convertibility of paper currency into gold was the primary policy objective. Sterilization of gold flows, though frequent, was more 'meeting the needs of trade' (passive monetary policy) than fighting unemployment (active monetary policy).
8. The gradual establishment of mint prices over time ensured that mint parities were in line with relative price levels; so countries joined the gold standard with exchange rates in equilibrium.
9. Current-account and capital-account imbalances tended to be offsetting for the core countries. A trade deficit induced a gold loss and a higher interest rate, attracting a capital inflow and reducing capital outflow. The capital-exporting core countries could stop a gold loss simply by reducing lending abroad.

13.1.7 Implications of Credible Commitment

Private parties reduced the need for balance-of- payments adjustment, via both gold-point arbitrage and stabilizing speculation. When the exchange rate was outside the spread, gold-point arbitrage quickly returned it to the spread. Within the spread, as the exchange value of a currency weakened, the exchange rate approaching the gold-export point, speculators had an ever-greater incentive to purchase domestic with foreign currency (a capital inflow). They believed that the exchange rate would move in the opposite direction, enabling reversal of their transaction at a profit. Similarly, a strengthened currency involved a capital outflow. The further the exchange rate moved toward a gold point, the greater the potential profit opportunity in betting on a reversal of direction; for there was a decreased distance to that gold point and an increased distance from the other point. This 'stabilizing speculation' increased the exchange value of depreciating currencies, and thus gold loss could be prevented. Absence of controls meant such private capital flows were highly responsive to exchange-rate changes.

13.1.8 Government Policies that Enhanced Stability

Specific government policies enhanced gold-standard stability. First, by the turn of the twentieth century, South Africa—the main world gold producer—was selling all its gold output in London, either to private parties or to the Bank of England. Thus the Bank had the means to replenish its gold reserves. Second, the orthodox-metallism ideology and the leadership of the Bank of England kept countries' monetary policies disciplined and in harmony. Third, the US Treasury and the central banks of the other core countries manipulated gold points, to stem gold outflow. The cost of exporting gold was artificially increased (for example, by increasing selling prices for bars and foreign coin) and/or the cost of importing gold artificially decreased (for example, by providing interest-free loans to gold importers).

Fourth, central-bank cooperation was forthcoming during financial crises. The precarious liquidity position of the Bank of England meant that it was more often the recipient than the provider of financial assistance. In crises, the Bank would obtain loans from other central banks, and the Bank of France would sometimes purchase sterling to support that currency. When needed, assistance went from the Bank of England to other central banks. Also, private bankers unhesitatingly made loans to central banks in difficulty.

Thus, 'virtuous' interactions were responsible for the stability of the gold standard. The credible commitment to convertibility of paper money at the established mint price, and therefore to fixed mint parities, was both a cause and an effect of the stable environment in which the gold standard operated—the stabilizing behavior of arbitrageurs and speculators, and the responsible policies of the authorities—and these three elements interacted positively among themselves.

13.1.9 Experience of Periphery

An important reason for periphery countries to join and maintain the gold standard was the fostering of access to core-countries' capital markets. Adherence to the gold standard connoted that the peripheral country would follow responsible macroeconomic policies and repay debt. This 'seal of approval', by reducing the risk premium, involved a lower interest rate on the country's bonds sold abroad, and very likely a higher volume of borrowing, thereby enhancing economic development.

However, periphery countries bore the brunt of the burden of adjustment of payments imbalances with the core (and other western European) countries. First, when the gold-exchange-standard periphery countries ran a surplus (deficit), they increased (decreased) their liquid balances in the United Kingdom (or other reserve-currency country) rather than withdraw gold from (lose gold to) the reserve-currency country. The monetary base of the periphery country increased (decreased), but that of the reserve-currency country remained unchanged. Therefore, changes in domestic variables—prices, incomes, interest rates, portfolios—that occurred to correct the imbalance were primarily in the periphery.

Second, when Bank Rate increased, London drew funds from France and Germany, which countries attracted funds from other European countries, which drew capital from the periphery. Also, it was easy for a core country to correct a deficit by reducing lending to, or bringing capital home from, the periphery. While the periphery was better off with access to capital, its welfare gain was reduced by the instability of capital import. Third, periphery-countries' exports were largely primary products, sensitive to world market conditions. This feature made adjustment in the periphery take the form more of real than financial correction.

The experience of adherence to the gold standard differed among periphery groups. The important British Dominions and colonies successfully maintained the gold standard. They paid the price of serving as an economic cushion to the Bank of England's financial situation; but, compared with the rest of the periphery, gained a stable long-term capital inflow. In southern Europe and Latin America, adherence to the gold standard was fragile. The commitment to convertibility lacked credibility, and resort to a paper standard occurred. Many of the reasons for credible commitment that applied to the core countries were absent. There were powerful inflationary interests, strong balance-of-payments shocks, and rudimentary banking sectors. The cost of adhering to the gold standard was apparent: loss of the ability to depreciate the currency to counter reductions in exports. Yet the gain, in terms of a steady capital inflow from the core countries, was not as stable or reliable as for the British Dominions and colonies.

13.1.10 Breakdown of Classical Gold Standard

The classical gold standard was at its height at the end of 1913, ironically just before it came to an end. The proximate cause of the breakdown of the classical gold standard was the First World War. However, it was the gold-exchange standard and the Bank of England's precarious liquidity position that were the underlying cause. With the outbreak of war, a run on sterling led Britain to impose extreme exchange control–a postponement of both domestic and international payments—making the international gold standard inoperative. Convertibility was not suspended legally; but moral suasion, legalistic action, and regulation had the same effect. The Bank of England commandeered gold imports and applied moral suasion to bankers and bullion brokers to restrict gold exports.

The other gold-standard countries undertook similar policies—the United States not until 1917, when it adopted extra-legal restrictions on convertibility and restricted gold exports. Commercial banks converted their notes and deposits only into currency. Currency inconvertibility made mint parities ineffective; floating exchange rates resulted.

13.1.11 Return to the Gold Standard

After the First World War, a general return to gold occurred; but the interwar gold standard differed institutionally from the classical gold standard. First, the new gold standard was led by the United States, not Britain. The US embargo on gold exports was removed in 1919, and currency convertibility at the pre-war mint price was restored in 1922. The gold value of the dollar rather than pound sterling was the typical reference point around which other currencies were aligned and stabilized. The core now had two center countries, the United Kingdom (which restored gold in 1925) and the United States.

Second, for many countries there was a time lag between stabilizing the currency in the foreign-exchange market (fixing the exchange rate or mint parity) and resuming currency convertibility. The interwar gold standard was at its height at the end of 1928, after all core countries were fully on the standard and before the Great Depression began. The only countries that never joined the interwar gold standard were the USSR, silver-standard countries (China, Hong Kong, Indochina, Persia, Eritrea), and some minor Asian and African countries.

Third, the 'contingency clause' of convertibility conversion, that required restoration of convertibility at the mint price that existed prior to the emergency (the First World War), was *broken* by various countries, and even core countries. While some countries (including the United States and United Kingdom) stabilized their currencies at the pre-war mint price, others (including France) established a gold content of their currency that was a fraction of the pre-war level: the currency was devalued in terms of gold, the mint price was higher than pre-war. Still others (including Germany) stabilized new currencies adopted after hyperinflation.

Fourth, the gold-coin standard, dominant in the classical period, was far less prevalent in the interwar period. All four core countries had been on coin in the classical gold standard; but only the United States was on coin interwar. The gold-bullion standard, non-existent pre-war, was adopted by the United Kingdom and France. Germany and most non-core countries were on a gold-exchange standard.

13.1.12 Instability of Interwar Gold Standard

The interwar gold standard was replete with forces making for *instability*.

1. The process of establishing fixed exchange rates was piecemeal and haphazard, resulting in disequilibrium exchange rates. Among core countries, the United Kingdom restored convertibility at the pre-war mint price without sufficient deflation, and had an over-valued currency of about ten per cent. France and Germany had undervalued currencies.
2. Wages and prices were less flexible than in the pre-war period.
3. Higher trade barriers than pre-war also restrained adjustment.
4. The gold-exchange standard economized on total world gold via the gold of the United Kingdom and United States in their reserves role for countries on the gold-exchange standard and also for countries on a coin or bullion standard that elected to hold part of their reserves in London or New York. However, the gold-exchange standard was unstable, with a conflict between (a) the expansion of sterling and dollar liabilities to foreign central banks, to expand world liquidity, and (b) the resulting deterioration in the reserve ratio of US and UK authorities.

   This instability was particularly severe, for several reasons. First, France was now a large official holder of sterling, and France was resentful of the United Kingdom. Second, many more countries were on the gold-exchange standard than pre-war. Third, the gold-exchange standard, associated with colonies in the classical period, was considered a system inferior to a coin standard.

5. In the classical period, London was the one dominant financial center; in the interwar period, it was joined by New York and, in the late 1920s, Paris. Private and official holdings of foreign currency could shift among the two or three centers, as interest-rate differentials and confidence levels changed.
6. There was maldistribution of gold. In 1928, official reserve-currency liabilities were much more concentrated than in 1913, British pounds accounting for 77% of world foreign-exchange reserves and French francs less than two per cent (versus 47 and 30% in 1913). Yet the United Kingdom held only seven percent of world official gold and France 13 per cent. France also possessed 39% of world official foreign exchange. The United States held 37% of world official gold.
7. Britain's financial position was even more precarious than in the classical period. In 1928, the gold and dollar reserves of the Bank of England covered only one-third of London's liquid liabilities to official foreigners, a ratio hardly greater than in 1913. UK liquid liabilities were concentrated on stronger countries (France, United States), whereas UK liquid assets were predominantly in weaker countries (Germany). There was ongoing tension with France, which resented the sterling-dominated gold-exchange standard and desired to cash in its sterling holding for gold, to aid its objective of achieving first-class financial status for Paris.
8. Internal balance was an important goal of policy, which hindered balance-of-payments adjustment, and monetary policy was influenced by domestic politics rather than geared to preservation of currency convertibility.
9. Credibility in authorities' commitment to the gold standard was not absolute. Convertibility risk and exchange risk could be high, and currency speculation could be destabilizing rather than stabilizing. When a country's currency approached or reached its gold-export point, speculators might anticipate that currency convertibility would not be maintained and that the currency would be devalued.
10. The 'rules of the game' were violated even more often than in the classical gold standard. Sterilization of gold inflows by the Bank of England can be viewed as an attempt to correct the overvalued pound by means of deflation. However, the US and French sterilization of their persistent gold inflows reflected exclusive concern

for the domestic economy and placed the burden of adjustment (deflation) on other countries.

11. The Bank of England did not provide a leadership role in any important way, and central-bank cooperation was insufficient to establish credibility in the commitment to currency convertibility. The Federal Reserve had three targets for its discount-rate policy: strengthen the pound, combat speculation in the New York stock market, and achieve internal balance—and the first target was of lowest priority. Although, for the sake of external balance, the Bank of England kept Bank Rate higher than internal considerations would dictate, it was understandably reluctant to abdicate Bank Rate policy entirely to the balance of payments, with little help from the Federal Reserve. To keep the pound strong, substantial international cooperation was required, but was not forthcoming.

13.1.13 Breakdown of Interwar Gold Standard

The Great Depression triggered the unravelling of the gold standard. The depression began in the periphery. Low export prices and debt-service requirements created insurmountable balance-of-payments difficulties for gold-standard commodity producers. However, US monetary policy was an important catalyst. In 1927 the Federal Reserve favored easy money, which supported foreign currencies but also fed the New York stock-market boom. Reversing policy to tame the boom, higher interest rates attracted monies to New York, weakening sterling in particular. The crash of October 1929, while helping sterling, was followed by the US depression.

This spread worldwide, with declines in US trade and lending. In 1929 and 1930 a number of periphery countries—both Dominions and Latin American countries—either formally suspended currency convertibility or restricted it so that currencies violated the gold-export point.

It was destabilizing speculation, emanating from lack of confidence in authorities' commitment to currency convertibility, which ended the interwar gold standard. In May 1931 there was a run on Austria's largest commercial bank, and the bank failed. The run spread to other eastern European countries and to Germany, where an important bank also collapsed. The countries' central banks lost substantial reserves; international financial assistance was too late; and in July 1931 Germany adopted exchange control, followed by Austria in October. These countries were definitively off the gold standard.

The Austrian and German experiences, as well as British budgetary and political difficulties, were among the factors that destroyed confidence in sterling, which occurred in mid-July 1931. Runs on sterling ensued, and the Bank of England lost much of its reserves. Loans from abroad were insufficient, and in any event taken as a sign of weakness. The gold standard was abandoned in September, and the pound quickly and sharply depreciated on the foreign-exchange market, as overvaluation of the pound would imply.

Following the UK abandonment of the gold standard, many countries followed, some to maintain their competitiveness via currency devaluation, others in response to destabilizing capital flows. The United States held on until 1933, when both domestic and foreign demands for gold, manifested in runs on US commercial banks, became intolerable. 'Gold bloc' countries (France, Belgium, Netherlands, Switzerland, Italy, Poland), with their currencies now overvalued and susceptible to destabilizing speculation, succumbed to the inevitable by the end of 1936.

The Great Depression was worsened by the gold standard: gold-standard countries hesitated to inflate their economies, for fear of suffering loss of gold and foreign-exchange reserves, and being forced to abandon convertibility or the gold parity. The gold standard involved 'golden fetters', which inhibited monetary and fiscal policy to fight the Depression. As countries left the gold standard, removal of monetary and fiscal policy from their 'gold fetters' enabled their use in expanding real output, providing the political will existed.

In contrast to the interwar gold standard, the classical gold standard functioned well because of a confluence of 'virtuous' interactions, involving government policies, credible commitment to the standard, private arbitrage and speculation, and fostering economic and political environment. We will not see its like again.

References

Bayoumi, T., B. Eichengreen, and M.P. Taylor, eds. 1996. *Modern perspectives on the gold standard*. Cambridge: Cambridge University Press.

Bordo, M.D., and F.E. Kydland. 1995.The gold standard as a rule: An essay in exploration. *Explorations in Economic History* 32: 423–464.

Bordo, M.D., and H. Rockoff. 1996. The gold standard as a 'good housekeeping seal of approval'. *Journal of Economic History* 56: 389–428.
Bordo, M.D., and A.J. Schwartz, eds. 1984. *A retrospective on the classical gold standard, 1821-1931*. Chicago: University of Chicago Press.
De Macedo, J.B., B. Eichengreen, and J. Reis, eds. 1996. *Currency convertibility: The gold standard and beyond*. London: Routledge.
Eichengreen, B. 1992. *Golden fetters: The gold standard and the Great Depression, 1919-1939*. New York: Oxford University Press.
Eichengreen, B., and M. Flandreau. 1997. *The gold standard in theory and history*. 2nd ed. London: Routledge.
Gallarotti, G.M. 1995. *The anatomy of an international monetary regime: The classical gold standard, 1880-1914*. New York: Oxford University Press.
Officer, L.H. 1996. *Between the dollar-sterling gold points*. Cambridge: Cambridge University Press.
Officer, L.H. 2001. Gold standard. In *EH.Net Encyclopedia*, website.

13.2 Silver Standard

The silver standard, the dominant monetary system for many centuries, lost much importance with the advent of the classical gold standard; and, due to US policy, residual monetary use of silver was virtually eliminated in the 1930s.

13.2.1 Definition of Silver Standard

A silver standard involves (a) a fixed silver content of the monetary unit, (b) 'free coinage' of silver, that is, privately owned silver in form other than domestic coin convertible into domestic silver coin at, or approximately at, the mint price (the inverse of the silver content of the monetary unit), (c) no restrictions on private parties (i) melting domestic coin into bullion, or (ii) importing or exporting silver in any form, and (d) full legal-tender status for domestic silver coin.

Other forms of money may exist, but silver is the primary money. Foreign silver coin may be given equal legal-tender status with domestic coin. Gold coin may be in circulation, but its value is in terms of the silver monetary unit and may fluctuate by weight, varying with the market gold-silver price ratio. Paper currency and deposits may exist, but, as liabilities of the issuer or bank, are payable in legal tender, that is, silver coin (or silver-convertible government or central-bank currency).

If silver (whether domestic or foreign coin, or both) constitutes the only money, then, even absent free coinage, the economy is clearly on a silver standard. This conclusion holds with gold coin circulating as well, providing it is circulating by weight or is a minor part of the money supply.

A silver standard might be effective even though the monetary system is legally bimetallic. If the coinage gold-silver price ratio is sufficiently below the market ratio, then gold, undervalued at the mint, will be sold on the world market (even in the form of melted domestic coin), while silver, overvalued, will be imported and coined. Ultimately, an effective silver standard may result.

Depreciation of the silver coinage involves an increased ratio of the legal (face) value of coins relative to silver content, usually by debasement (reducing the silver content, whether weight or fineness, of given-denomination coins) rather than by increasing the denomination of existing (given-weight-and-fineness) coins. In England, the penny (of sterling, 11/12th fineness) was steadily reduced in size from 24 grains in the eighth century to less than 1/3 that weight in 1601.

A silver standard, just as the gold standard, provides a constraint on the money stock. Depreciation of silver coinage was a way of escaping that constraint, even though the authority's objective typically was to increase government revenue (in the form of seigniorage) and/or to change the coinage ratio (under legal bimetallism).

13.2.2 Countries on Silver Standard to 1870

A silver standard first occurred in ancient Greece. Notwithstanding generally legal bimetallism, silver was everywhere the effective metallic standard—or at least the far-more-important coined metal in the money stock—well into the eighteenth century. Because of its relative scarcity and high density, gold was always much more valuable than silver on a per-ounce basis: coinage and market ratios were far above unity. So, with most transactions of low value compared with the unit of account, silver was better suited than gold to serve as a medium of exchange. In US history, 'one dollar' was both the smallest gold piece and the largest silver piece ever coined.

In England, from the Anglo-Saxon period until the late thirteenth century, the only coin in existence (with rare exceptions) was the silver penny, with 240 pence coined ideally from one pound of silver and later constituting one pound sterling (where ‘sterling,’ of course, denotes silver). This was a silver standard by default. With coinage of gold, in 1257, there was legal bimetallism; but the practice of denominating gold coins in (silver) shillings and pence was implicit recognition of an effective silver standard. Even the popular, consistently coined, (gold) guinea, first issued in 1663, was left to find its own market value in shillings and pence. However, by the turn of the eighteenth century, foreign gold-silver price ratios had been falling and, having been increased greatly in 1696, the British coinage ratio was not subsequently reduced enough to compensate. England went briefly on a bimetallic standard, and then on an effective gold standard, legalized in 1774 and 1816.

In the United States, since colonial times a silver standard was in effect, based on the Spanish dollar, the primary circulating silver coin, which varied much in weight and fineness. Yet the dollar was accepted everywhere at face-value in terms of local (individual-state) pound-shilling-pence units of account. Gold coins were rated in dollars according to fine-metal content. The Coinage Act of 1792 placed the United States on a legal bimetallic standard; but the coinage ratio soon fell below the (increasing) world-market ratio. An effective silver standard resulted, until the coinage ratio was corrected in 1834.

In 1870, just before Germany united and established the gold standard (using as financing the French indemnity, emanating from the Franco-Prussian War), Netherlands, Denmark, Norway, Sweden, India, China, Straits Settlements, Hong Kong, Dutch East Indies, Mexico and some German states were on a silver standard. In the 1870s these European countries (and Dutch East Indies) abandoned silver in favor of gold. By 1885 almost all of western Europe—along with the United States, Britain, its Dominions and various colonies—was on gold.

13.2.3 Asian Abandonment of Silver Standard Prior to World War I

Traditionally, Asian countries preferred silver to gold for both monetary and nonmonetary use, and the low market ratios in the Far East reflected that fact. The silver standard continued after 1885 in the Asian countries listed above. Further, in the 1880s the Philippines and Japan went on de facto silver.

Until 1873, bimetallic France kept the world-market gold-silver price ratio around a narrow band centered on the French coinage ratio of 15½.

When France ended bimetallism in 1873, the market ratio lost its anchor and escalated tremendously. The exchange rates between silver-standard and gold-standard currencies also lost their anchor. Following the market gold-silver price ratio, silver currencies depreciated greatly with respect to gold currencies. Exports were enhanced, imports were more expensive, debt and other obligations stated in terms of gold or gold currencies increased greatly in domestic currency, domestic inflation increased, and foreign investment was discouraged due to exchange-rate instability.

The problem of a depreciating currency was especially acute for India, which had the obligation of substantial recurring sterling-denominated 'home charges' to Britain (for debt service, pensions, military and other equipment, and so forth). In 1893 India abandoned the silver standard, and in 1898 went on the gold-exchange standard, pegging the (silver) rupee against the pound sterling.

In 1897 Japan switched from a de facto silver standard under legal bimetallism to a monometallic gold-coin standard, using as financing the indemnity received from defeated China in the Sino-Japanese War. In 1903 the Philippines adopted a gold-exchange standard, with the (silver) peso pegged to the US (gold) dollar. The impetus was transfer of the country from Spain to the United States, thanks to US victory in the Spanish-American War.

Mexico, a large silver producer, with both commodity exporters and silver producers in favor of a continued silver standard, finally adopted a gold-coin standard in 1905. At the beginning of the First World War, the silver standard encompassed only China, Hong Kong, and a few minor countries.

13.2.4 Termination of Silver Standard

The final blow to the silver standard was delivered by the United States, ironically after it left the gold standard. In December 1933, when the (fluctuating) market price of silver was 44 cents per ounce, President Roosevelt proclaimed that US mints should purchase all new domestically produced silver at a net price (to the depositor, or seller) of 64.65 cents per ounce (half the official, but inoperative, mint price of silver). In 1934 this policy was reinforced by the Silver Purchase Act, which directed the Treasury to purchase silver at home and abroad as long as (a) the Treasury stock of gold constituted less than one-quarter its total monetary stock, and (b) the market price did not exceed the US official mint price. Subsequently, the president ordered that all silver (with minor exceptions) then situated in the continental United States was to be delivered to US mints, at a net price of 50.01 cents per ounce. In 1935, in response to a higher foreign market price of silver (largely due to the US silver-purchase policy itself!), the president increased the net price for newly produced domestic silver to 71.11 cents.

The reason for the US silver-purchase policy was to provide a subsidy to the (politically powerful) domestic silver producers. Inadvertently, the policy effectively destroyed what remained of the silver standard. The last major country on the silver standard was China. As the gold-standard world suffered monetary and real deflation in 1929–30, the price of silver fell. The Chinese, silver-based, currency (yuan) therefore depreciated against the, gold-based, currencies of important trading partners (Britain, India, Japan). The enhanced competitiveness of export and import-competing industries, and resulting balance-of-payments surplus, prevented deflation. China lost some 'silver protection' in 1931, after Britain, India and Japan left the gold standard, as the yuan appreciated against the pound, rupee and yen; but the United States was still on the gold standard, and the yuan continued to fall, slightly, against the dollar. After the United States abandoned the gold standard, in 1933, the yuan appreciated against all four currencies.

While China had lost its 'silver protection' from the world depression, it nevertheless retained the silver standard and probably suffered less economically than its main trading partners. Disaster struck with the US silver policy of 1933–1934. The huge increase in the US and market price of silver involved a corresponding appreciation of the yuan. Loss of competitiveness, balance-of-payments deficit, export of silver (and gold) to finance the deficit, and deflation followed. China had no choice but to leave the silver standard, effectively in 1934, and legally in 1935.

Other silver-standard, as well as silver-using, countries were also adversely affected by the US policy. Hong Kong followed China, and left the gold standard in 1935. Though not on the silver standard, various Latin American countries had a large silver coinage. These were token coins (face-value higher than metallic-content value). Nevertheless, the high US price for silver encouraged the melting and export of these coins. The affected countries resorted to debasement and re-coining in order to retain their silver coinage.

Mexico was a special case. Silver coins constituted a high proportion of its money supply; but, as the world's largest producer of silver, Mexico benefited from a higher price for a major export. However, as other countries left the silver standard, the price of silver began to fall, and this advantage was reduced. Mexico prohibited melting or export of silver coins in 1935, and replaced the coins with paper money. Later, re-coinage occurred, and melting and export were again permitted. Yet the damage had been done, and Mexico was now on a 'managed paper standard', having lost the discipline provided by metallic money. In sum, in the 1930s, a US domestic-oriented policy reduced considerably such monetary use of silver as remained.

References

Bojanic, A.N. 1994. *The silver standard in two late adherents: Mexico and India*. Ph.D. thesis, Auburn University.

Brandt, L., and T.J. Sargent. 1989. Interpreting new evidence about China and U.S. silver purchases. *Journal of Monetary Economics* 23: 31–51.

der Eng, P. V. 1993. *The silver standard and Asia s integration into the world economy, 1850-1914.* Canberra: Australian National University.

Eichengreen, B., and M. Flandreau. 1996. The geography of the gold standard. In *Currency convertibility: The gold standard and beyond*, ed. J.B. de Macedo, B. Eichengreen, and J. Reis. London and New York: Routledge.

Einzig, P. 1970. *The history of foreign exchange*, 2nd ed. London: Macmillan.

Feavearyear, Sir A. 1963. *The pound sterling: A history of English money,* 2nd ed. Oxford: Clarendon Press.

Friedman, M. 1992. *Money Mischief: Episodes in monetary history.* New York: Harcourt Brace Jovanovich.

Lai, C., and J.J. Gau. 2003. The Chinese silver standard economy and the 1929 Great Depression. *Australian Economic History Review* 43: 155–168.

Leavens, D.H. 1939. *Silver money.* Bloomington, Indiana: Principia Press.

Officer, L.H. 1996. *Between the dollar-sterling gold points.* Cambridge: Cambridge University Press.

Redish, A. 2000. *Bimetallism: An economic and historical analysis.* Cambridge: Cambridge University Press.

Wilson, T. 2000. *Battles for the standard: Bimetallism and the spread of the gold standard in the nineteenth century.* Aldershot: Ashgate.

13.3 Bimetallism

A bimetallic monetary standard is a combination of two metallic standards, each of which could in principle stand alone, and often evolved into de facto monometallism.

13.3.1 The Nature of Bimetallism

Bimetallic metals are usually gold and silver, but there are exceptions. Ancient Rome was temporarily on a silver-bronze standard; in the eighteenth century, Sweden and Russia experienced a silver-copper standard.

Under bimetallism, both gold and silver coins are full legal tender. The unit of account (dollar, franc, and so on) is defined in terms of a fixed weight both of pure gold and of pure silver. So there is a fixed legal (mint, coinage) gold-silver price ratio: number of grains or ounces of silver per grain or ounce of gold. Both gold and silver enjoy free coinage (the government prepared to coin bars of either metal deposited by any party) and are full-bodied (have legal or face-value equal to metallic value). Token subsidiary (always silver) coins can exist. Subsidiary coins are fractions of (have face-value less than) the unit of account; token coins have face-value less than metallic (inherent) value, and invariably have restricted legal-tender power. Token coins were not adopted by bimetallic countries until late in their experience with bimetallism, and in conjunction with the process of terminating that standard.

Private parties may melt, import, and export coins (domestic or foreign) of either metal. There is no restriction on non-monetary uses of the monetary metals. Paper currency and deposits may exist; they are convertible into legal-tender coins, either directly or via government-issued paper currency (itself directly convertible into coin). Both private parties and the government may choose the metallic coin, or mixture of coins, in which to discharge debt (including paper currency). However, a private party does not have the right to a direct governmental exchange of gold for silver, or silver for gold. Logically, though, domestic gold and silver coin would exchange privately at the mint ratio.

13.3.2 Advantages and Disadvantages of Bimetallism

Bimetallism has four advantages. First, it embodies two sets of coins–one from a metal with a high value-weight ratio (gold), the other from a metal with a low ratio (silver). These provide a medium of exchange for a wide range of economic transactions. The range can be extended in both directions: upper, via paper currency and deposits; lower, via token subsidiary coins. Neither is incompatible with a bimetallic standard. Second, as does a monometallic standard, the bimetallic standard provides a constraint on the money supply and therefore on inflation; for the legal-tender coins constitute the monetary base (given government-issued legal-tender paper, perhaps the 'super monetary base'), and the government must acquire one or the other metal to increase the base. Because there is coinage on demand, there is also a check on reduction to the monetary base, and on deflation. Third, a bimetallic country or bloc of countries accommodates shocks, so that resulting effects on monometallic-countries' money supplies are dampened. This is done by stabilizing the gold-silver price ratio ('market ratio') on the world market, the bullion market, where non-monetary gold and silver (generally bars) are traded either among themselves or individually for some important currency. Fourth, in stabilizing the market gold-silver price ratio, the bimetallic country or bloc also stabilizes the exchange rates between 'gold currencies' and 'silver currencies'. Otherwise, these exchange rates would fluctuate, defeating one of the usual purposes of metallic standards.

The alleged disadvantage of bimetallism (relative to monometallism) is that it is unstable. Suppose the bimetallic-country's mint ratio initially is in the neighborhood of the market ratio. A shock in the world supply of one metal can change the market ratio so that the mint ratio is now outside its neighborhood. If the resulting market ratio is above (below) the mint ratio, then silver (gold) is 'bad' money, overvalued at the mint; domestic payments will tend to be made in that, relatively cheaper, coin rather than gold (silver), the 'good' money, undervalued at the mint and relatively expensive in the market. Good money will tend to be exported to settle balance-of-payments surpluses, bad money imported to finance balance-of-payments deficits. If the divergence between the market and mint ratio is large, 'bimetallic arbitrage' occurs, whereby good money is melted and traded on the bullion market for the bad metal, and the bad metal imported to be coined. In both situations, Gresham's law is operative: bad money drives out good.

Given sustained payments imbalances and/or a large and persistent divergence between the market and mint ratio, bad-money monometallism results. (The good money may be eliminated from the money supply, or circulate at a market-determined value—available only at a premium.) To avoid this, the mint ratio could be altered to remain in conformity with the market ratio. If the mint ratio is under-corrected, monometallism is not stemmed; if the mint ratio is over-corrected, monometallism in the opposite metal can occur. Successive changes in the market ratio can lead to alternating effective gold monometallism and silver monometallism, under the rubric of legal bimetallism. There are costs to such an alternating monetary standard; there are also costs in periodically altering the mint ratio.

13.3.3 Theories of Bimetallic Stabilization

Stabilizing bimetallic arbitrage happens as follows. Suppose a shock occurs, new gold discoveries, that decrease the market ratio: the market price of non-monetary gold falls relative to silver. The market ratio now is below the mint ratio, so gold is 'bad' (overvalued) and silver 'good' (undervalued) money. Silver leaves the monetary system to be sold in the world (bullion) market, with gold purchased with the proceeds and coined. First, the arbitrageurs make a profit: the value of the gold coins they obtain is greater than the value of the silver coins they initially sold. Second, there is increased supply of silver (the appreciated metal) and increased demand for gold (the depreciated metal) in the bullion market—the two transactions constituting one arbitrage transaction. The result is an increase in the market ratio, which rises toward the mint ratio. Thus, the incentive for the arbitrage is eliminated. Third, the composition of the money supply of the bimetallic country changed, with a higher proportion of gold to silver. The bimetallic country stabilized the market ratio (and incidentally the exchange rates between gold and silver currencies), via the endogenous gold-silver composition of its money supply.

This mechanism is effective only to the extent that the bimetallic country has sufficient stock of the undervalued metal to return the market ratio close to the mint ratio, so that the incentive to arbitrage vanishes before monometallism in the overvalued metal results. However, the situation is not so dire, because costs of arbitrage imply 'gold-silver price-ratio' points that define a band for the market ratio within which the ratio can fluctuate without triggering bimetallic arbitrage. If the bimetallic-country's commitment to its mint ratio is absolutely credible, then stabilizing speculation exists within the bimetallic-arbitrage band, such that the market ratio turns away from its nearest bound and towards the mint ratio. The situation is analogous to stabilizing speculation within gold-point spreads, under the international gold standard.

Two other forces making for bimetallic stability have been suggested by Marc Flandreau. The first is 'metal-specific arbitrage' between the bullion and monetary markets. If a metal depreciates on the bullion market by more than coinage and associated costs, then owners of bars in that metal will coin them in lieu of supplying them to the bullion market. If a metal appreciates by more than melting and associated costs of bringing that coined metal to the market, then holders of coin of that metal will melt them and supply them to the market. The reduced supply of the depreciated metal and increased supply of the appreciated metal act to return the market ratio towards the mint ratio. Unlike bimetallic arbitrage, these are independent transactions. Therefore the costs of metal-specific arbitrage are below the costs of bimetallic arbitrage, and the former provide a 'metal-specific band' located within the 'bimetallic arbitrage band.' So metal-specific arbitrage is a stabilizing mechanism that becomes operative before bimetallic arbitrage.

The second force involves the bimetallic country (France) transacting with a gold-currency country (England) and a silver-currency country (Germany). There are franc-sterling gold points, and franc-mark silver points. Expressing exchange rates as percentage deviations from parity and specie points in percentage terms, the franc/sterling - franc/mark exchange-rate differential (via triangular arbitrage) proxies the mark/sterling exchange rate. Also, implicit mark-sterling parity (via franc bilateral parities) corresponds to the mint ratio. On the assumption of no bilateral specie-point violations, the mark-sterling exchange rate has as upper (lower) bound the sum (negative *sum*) of the franc-sterling export (import) point and the franc-mark import (export) point. Now, the mark-sterling exchange rate is itself a good representation of the gold-silver market price ratio, because the Bank of England (Bank of Hamburg) supports, within a narrow band, a fixed sterling (mark) price of gold (silver). For the market ratio above the mint ratio (parity), so that silver is overvalued, the upper bound correctly involves exporting gold (sterling) and importing silver (marks). The gold-silver market price ratio has a bimetallic-arbitrage band that is approximately double the width of the franc-sterling and franc-mark bilateral specie-point spreads. Hence specie flows to settle and adjust payments imbalances occur prior to bimetallic arbitrage.

Suppose that a bimetallic country has lost all its undervalued ('good') metal, so it has become monometallic in its overvalued coinage. Nevertheless, Oppers (2000) shows that a bimetallic-arbitrage band could exist, given that there is a second bimetallic country with a different mint ratio. The two-countries' mint ratios each constitute a bound to the market ratio, with, as usual, a market ratio beyond a bound giving rise to arbitrage that returns the market ratio to the band. For this mechanism to operate, both countries must actually or potentially have large amounts of both coined metals in their money stock, where 'large' means relative to shocks in the bullion market.

13.3.4 Bimetallism Prior to the 19th Century

The Persian Empire had the first bimetallic standard, with a mint ratio of 13½ to 1 (all known mint ratios are in favor of gold) for a long time. This ratio undervalued silver relative to the ratio elsewhere, and presumably merchants took advantage of the price-ratio discrepancies in their regular dealings. The Roman Empire was often gold-silver bimetallic, but periodically debased the coinage. The likely reason was to increase seigniorage rather than to realign the mint ratio in conformity with the market ratio or the mint ratio in other lands. Until the mid-nineteenth century, bimetallism was the legal standard in Europe (including England), though the mint ratio was often altered. Traditionally, the gold-silver price ratio was lower in China and India than in Europe.

England was legally on a bimetallic standard from the mid-thirteenth century, when gold was first coined. The mint ratio was often changed. England was effectively on a silver standard until late in the seventeenth century, because the British mint ratio was generally below European gold-silver price ratios. Gold coins passed at a market price (in terms of the silver shilling) rather than face-value, again indicative of a silver standard. In 1663 the (gold) guinea was coined, with a legal value of 20 (silver) shillings. The silver coins in circulation were in horrible condition, due in part to past debasement, in part to private clipping and sweating of the coins. So the market price of the guinea increased above 20 shillings—to as much as 30 shillings—implying a gold-silver price ratio that effectively overvalued gold relative to Continental ratios. England was in process of switching from an effective silver to an effective gold standard.

In 1696 silver was recoined, so the coins became full-bodied again, and a ceiling (periodically reduced) was placed on the market price of the guinea. The result was that, for a brief period at the turn of the eighteenth century, England had effective bimetallism, with full-bodied coins of both metals in circulation. However, gold continued to be overvalued and silver undervalued; silver was exported, gold imported; and a de facto gold standard resulted. It became a de jure standard, via legislations restricting the legal-tender power of silver (1774) and effectively ending free coinage of silver (1816).

The Coinage Act of 1792 placed the United States on a legal bimetallic standard. The mint ratio (15 to 1)—selected because it was approximately the market ratio at the time—turned out to overvalue silver, because the market ratio increased. By 1823 gold had virtually gone from circulation, and an effective silver standard resulted. In 1834 Congress increased the ratio to 16.0022 (in 1837, revised slightly, to 15.9884). From 1834 to 1873, the world gold-silver price ratio was consistently below 16, so the new ratio overvalued gold, and an effective gold standard resulted. However, the export of full-bodied Mexican (silver) dollars and US subsidiary silver protected the circulation of underweight foreign silver pieces, which circulated at face-value; so in a sense effective bimetallism continued. Only in the early 1850s, when the market gold-silver price ratio fell (due to gold discoveries and new production), did the United States begin to lose its remaining silver coins. In 1853, to retain the silver, Congress reduced subsidiary coins (below a dollar) to token status, with limited legal-tender power. The United States now was on a de facto gold standard. Legal bimetallism remained until 1873, when coinage of the silver dollar was terminated. One year later, silver was virtually demonetized; all silver coins (including the dollar) were restricted to maximum legal tender of five dollars in any payment.

13.3.5 Bimetallic France in the 19th Century

In 1803 France made the franc the monetary unit, and solidified and made effective the mint ratio of 15½ that had been established in 1785. From the end of the Napoleonic Wars until 1873, while France retained that bimetallism, the market gold-silver price ratio remained in the neighborhood of 15½. (Also, exchange rates among gold, silver, and bimetallic countries were stable.) The stability of the market ratio was remarkable, in the face of severe shocks to the bullion market. In the 1850s gold production increased tremendously due to gold discoveries in California and Australia, putting strong downward pressure on the market price ratio. In the 1860s gold production stopped increasing, and exploitation of Nevada silver discoveries put strong upward pressure on the ratio.

The steady market gold-silver price ratio was due primarily to the continued bimetallism of France, which acted as a buffer to shocks and thus stabilized the gold-silver market price ratio. What gave France this power was its large economic size, the substantial amounts of both gold and silver coins in its circulation, and its credible commitment to bimetallism at an unchanged mint ratio. Therefore, French bimetallic arbitrage operated—in the 1850s and early 1860s via gold imported and coined and silver melted and exported, in the later 1860s via the opposite activities. Stabilizing speculation within the bimetallic-arbitrage band, stabilizing bilateral specie flows, and metal-specific arbitrage were also elements in the French stabilization service. In 1865 the French stabilizing force was enhanced by formation of the Latin Monetary Union (LMU), in which France, Belgium, Switzerland, and Italy adopted a common bimetallism.

Some scholars, especially Oppers (1995, 2000), believe, rather, that France underwent serial monometallism, with bimetallism transformed to a de facto silver standard in the 1830s and 1840s, and the latter yielding to a de facto gold standard in the 1860s. Yet a parity band (with stabilizing speculation within the band) existed, with the French mint ratio the lower bound and the US mint ratio the upper bound in 1834–1861, followed subsequently by the French ratio the upper bound and the Russian ratio the lower bound. This interpretation of history is doubtful, for the strong propensity to use both metallic currencies was characteristic only of France. Also, Russia's mint ratio was inoperative at the time, as the country had an inconvertible paper currency.

In the early 1860s the future LMU countries, if not on a de facto gold standard, were certainly moving towards it. With the market ratio below the mint ratio, silver was being lost. To protect silver circulation, the individual countries made subsidiary coins token currency; while in 1866 the LMU came into effect, mandating reduction of the silver content and restriction of the legal-tender power of all silver coins except the largest, that is, the five-franc piece, which remained full-bodied.

French, LMU, and world bimetallism ended in the 1870s. The proximate cause was Germany's move to a gold standard, financed by the

French indemnity that resulted from the Franco-Prussian War. Germany's release of silver put upward pressure on the gold-silver market price ratio. France was not prepared to accept the gold loss and silver inflow that would result from continued adherence to bimetallism. France (and Belgium) limited silver coinage in 1873, followed by the LMU mandating limits on coinage of the five-franc silver piece in 1874–1876. In 1878 coinage of that piece was terminated. The existing five-franc coins retained full legal-tender power. France, along with Belgium and Switzerland, went on a 'limping' gold standard, redeeming government-issued paper money in either gold or silver at the discretion of the authority.

References

Eichengreen, B., and M.R. Flandreau. 1996. Blocs, zones and bands: International monetary history in light of recent theoretical developments. *Scottish Journal of Political Economy* 43: 398–418.

Einzig, P. 1970. *The history of foreign exchange,* 2nd ed. London: Macmillan.

Flandreau, M.R. 1997. As good as gold? Bimetallism in equilibrium. In *Monetary standards and exchange rates,* ed. M.C. Marcuzzo, L.H. Officer, and A. Rosselli. London: Routledge.

Flandreau, M.R. 2002. 'Water seeks a level': Modeling bimetallic exchange rates and the bimetallic band. *Journal of Money, Credit and Banking* 34: 491–519.

Flandreau, M.R. 2004. *The glitter of gold*. Oxford: Oxford University Press.

Friedman, M. 1990. Bimetallism revisited. *Journal of Economic Perspectives* 4 (4): 85–104.

Gallarotti, G.M. 1994. The scramble for gold: Monetary regime transformation in the 1870s. In *Monetary regimes in transformation,* ed. M.D. Bordo and F. Capie. Cambridge: Cambridge University Press.

Leavens, D.H. 1939. *Silver money.* Bloomington, Indiana: Principia Press.

Martin, D.A 1968. Bimetallism in the United States before 1850. *Journal of Political Economy* 76: 428–442.

Martin, D.A. 1973. 1853: The end of bimetallism in the United States. *Journal of Economic History* 33: 825–844.

Officer, L.H. 1996. *Between the dollar-sterling gold points.* Cambridge University Press.

Oppers, S.E. 1995. Recent developments in bimetallic theory. In *International monetary systems in historical perspective*, ed. J. Reis. Houndmills/London: Macmillan.

Oppers, S.E. 2000. A model of the bimetallic system. *Journal of Monetary Economics* 46: 517–533.

Redish, A. 1994. The Latin Monetary Union and the emergence of the international gold standard. In *Monetary regimes in transformation,* ed. M.D. Bordo and F. Capie. Cambridge: Cambridge University Press.

Redish, A. 2000. *Bimetallism: An economic and historical analysis.* Cambridge: Cambridge University Press.

Velde, F.R., and W.E. Weber. 2000. A model of bimetallism. *Journal of Political Economy* 108: 1210–1234.

Wilson, T. 2000. *Battles for the standard*. Aldershot: Ashgate.

CHAPTER 14

Classical Gold Standard

14.1 International Monetary Regimes: The Gold Standard

Originally published in *Handbook of the History of Money and Currency*, edited by Stefano Battilossi, Youssef Cassis, and Kazuhiko Yago. Springer Nature, 2020, pp. 599–631.

14.1.1 Introduction

The classical gold standard is the most famous monetary system that ever existed, with its heyday lasting a third of a century. By the time World War I began, the gold standard had become the predominant national and international monetary system in the world. Countries may be allocated to different groups, depending on the importance of the country to the working of the gold standard, the type of gold standard to which the country adhered, and the extent to which the country observed the standard. Whether automatic or policy-induced, there are implications for the money supply. The main theme is that the gold standard exhibited both elements that promoted stability and forces that fostered instability. Modern time-series analysis has been used to examine various facets of the gold standard, especially the roles of the core countries (Britain, France, Germany, and the United States). While there is apparent consensus on

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_14

some aspects of the gold standard, controversies continue, and there remains room for further research and reflection.

14.1.2 Countries on Gold Standard

14.1.2.1 Legal Versus Effective Monetary Standard

Countries effectively on the gold standard and the periods during which they were on gold are listed in Table 14.1. The effective monetary standard of a country is distinguished from its legal standard. For example, a country legally on bimetallism usually was on either an effective gold or effective silver monometallic standard, depending on whether the country's "mint-price ratio" (the ratio of its mint price of gold to mint price of silver) was greater or less than the world price ratio. In contrast, a country might be legally on a gold standard, but its banks (and government) have "suspended specie (gold) payments" (refusing to convert their notes into gold), so that the country is in fact on a "paper standard."

Table 14.1 strives to incorporate all time periods (and only time periods) when a country was on an operational, or "effective," gold standard, irrespective of the legal standard; but in some cases only beginning and ending dates on gold can be discerned. The criterion adopted is that a country is deemed on the gold standard if (1) gold was the predominant effective metallic money, (2) specie payments were in force, and (3) there was a limitation on the coinage and/or the legal-tender status of silver (the only practical and historical competitor to gold), thus providing institutional or legal support for the effective gold standard emanating from (1) and (2).

The years 1880–1913 are generally construed as "the heyday of the gold standard," because throughout this period the "core countries" (Britain, France, Germany, United States), along with Scandinavia and several Western European countries, were continuously on gold. To quote (Flandreau et al. 1998, p. 150): "The big players were on gold and this is why economic history, rightly, puts the dates 1880–1913 on the gold standard."

In 1870 only Britain, two of its dependencies (Australia, Canada) and two countries also closely aligned with Britain economically and politically (Argentina and Portugal) were on the gold standard. Of all other countries, only the United States had ever been on an effective gold standard. By 1900, and even more so by 1914 (ironically, just before the gold standard collapsed, with World War I), almost every economically important country in the world had adopted gold. How did gold monometallism

Table 14.1 Countries on gold standard

Country	*Type of standard*	*Period*
Core countries		
Center country		
Britain	Coin	1774–1797, 1821–1914
Other core countries		
United States	Coin	1834–1861, 1879–1917
France	Coin	1878–1914
Germany	Coin	1871–1914
Inner Periphery		
British Colonies and Dominions		
Australia	Coin	1852–1915
Canada	Coin	1854–1914
Ceylon	Coin	1901–1914
India	Exchange (British pound)	1898–1914
Western Europe		
Austria-Hungary	Coin	1892–1914
Belgium	Coin	1878–1914
Italy	Coin	1884–1894
Liechtenstein	Coin	1898–1914
Netherlands	Coin	1875–1914
Portugal	Coin	1854–1891
Switzerland	Coin	1878–1914
Scandinavia		
Denmark	Coin	1872–1914
Finland	Coin	1877–1914
Norway	Coin	1875–1914
Sweden	Coin	1873–1914
Outer Periphery		
Eastern Europe		
Bulgaria	Coin	1906–1912
Greece	Coin	1885, 1910–1914
Montenegro	Coin	1911–1914
Romania	Coin	1890–1914
Russia	Coin	1897–1914
Middle East		
Egypt	Coin	1885–1914

(continued)

Table 14.1 (continued)

Country	*Type of standard*	*Period*
Turkey (Ottoman Empire)	Coin	1881–1914
Asia		
Japan	Coin	1897–1917
Philippines	Exchange (US dollar)	1903–1914
Siam	Exchange (British pound)	1908–1914
Straits Settlements	Exchange (British pound)	1906–1914
Mexico and Central America		
Costa Rica	Coin	1896–1914
Mexico	Coin	1905–1913
South America		
Argentina	Coin	1867–1876, 1883–1885, 1900–1914
Bolivia	Coin	1908–1914
Brazil	Coin	1888–1889, 1906–1914
Chile	Coin	1895–1898
Ecuador	Coin	1898–1914
Peru	Coin	1901–1914
Uruguay	Coin	1876–1914
Africa		
Eritrea	Exchange (Italian lira)	1890–1914
German East Africa	Exchange (German mark)	1885–1914
Italian Somaliland	Exchange (Italian lira)	1889–1914

Source Bulgaria—Dimitrova and Fantacci (2010, pp. 190, 194). Korea and Taiwan—Conant (1915, pp. 566–568). Other countries—Officer (2008, Table 1)

Britain includes colonies (except British Honduras) and possessions without a national currency: New Zealand and certain other Oceanic colonies, South Africa, Guernsey, Jersey, Malta, Gibralter, Cyprus, Bermuda, British West Indies, British Guiana, British Somaliland, Falkland Islands, other South and West African colonies. Britain first limited legal tender of silver in 1774, terminated free coinage of silver in 1798.

Table 14.1 (continued)

For precise dates and internal geographic exceptions for US period on gold standard, see Officer (1996, pp. 16–17). United States includes countries and territories with US dollar as exclusive or predominant currency: British Honduras (from 1894), Cuba (from 1898), Dominican Republic (from 1901), Panama (from 1904), Puerto Rico (from 1900), Alaska, Aleutian Islands, Hawaii, Midway Islands (from 1898), Wake Island, Guam, and American Samoa.

France includes Tunisia (from 1891) and all other colonies except Indochina.

Canada includes Newfoundland (from 1895). India includes British East Africa, Uganda, Zanzibar, Mauritius, and Ceylon (to 1901). Austria-Hungary includes Montenegro (to 1911). Belgium includes Belgian Congo. Netherlands includes Netherlands East Indies. Portugal includes colonies, except Portuguese India. Denmark includes Greenland and Iceland. Japan includes Korea and Taiwan (both from 1904). Straits Settlements includes Borneo. German East Africa and Italian Somaliland; beginning dates are approximate.

For other gold-standard lists, see Bordo and Schwartz (1996, pp. 20–22), Meissner (2005, p. 391), Martin-Aceña (2007, pp. 97–100), Mitchener and Weidenmier (2015, pp. 486, 508).

achieve its primacy? And, in particular, what explains the "scramble for gold" (or "rush to gold") that began in the 1870s?

14.1.2.2 Center Country

The country grouping in Table 14.1 reflects the importance of countries to establishment and maintenance of the standard. Consider a "core country" as a country of high importance to that end. Then Britain was the "center country," and thus the most important core country. It was the earliest country on a gold standard and was indispensable to the spread and functioning of the gold standard. "London was the center for the world's principal gold, commodities and capital markets… [There were] extensive outstanding sterling-denominated assets, and… many countries substituted sterling for gold as an international reserve currency" (Bordo 1993, p. 162).

For centuries, Britain had been on an effective silver standard under legal bimetallism. The country switched to an effective gold standard early in the eighteenth century, solidified by the (mistakenly) gold-overvalued mint-price ratio established by Isaac Newton, Master of the Mint, in 1717. In 1774 the legal-tender property of silver was restricted, and Britain entered the gold standard in the full sense. In 1798 coining of silver was suspended, and in 1816 the gold standard was formally adopted, ironically during a paper-standard regime (the "Bank Restriction Period," of 1797–1821), with the gold standard effectively resuming in 1821 and remaining until 1914.

14.1.2.3 Other Core Countries

Lindert identifies the pound sterling, French franc, and German mark as "key currencies," the most important reserve currencies. "The role of world banker was performed by Britain, France, and Germany in these years [1900–1913] on a scale unmatched either before or since" (Lindert 1969, p. 1). Flandreau and Jobst (2005), using the criterion of international circulation of domestic currencies (measured by number of recorded geographic exchange-market quotations), also place these three currencies in the top tier. The core countries Germany and France switched from bimetallism and silver to gold in 1871 and 1878, respectively.

It is controversial whether the United States should also be considered a core country. If a large circulation of gold coin is the criterion (Gallarotti 1995, p. 23), then the United States belongs in the group. If the existence of a central bank is required, then the United States does not so belong—a judgment also reached according to the Flandreau-Jobst criterion. However, tipping the scales in favor of inclusion is the fact that the United States was a heavyweight in the world economy, with large shares of would output, trade, and investment. Tullio and Wolters (2000, p. 62) state bluntly: "by 1910 US real GDP was three times UK GDP." Indeed, most scholars show revealed preference for inclusion, because their "heyday of the gold standard" begins only after the United States returned to the gold standard in 1879, thus completing the core group.

The United States was on an effective silver standard dating back to colonial times, legally bimetallic from 1786, and on an effective gold standard from 1834. The legal gold standard began in 1873–1874, when Congressional Acts ended silver-dollar coinage and limited legal tender of existing silver coins. Ironically and again, the move from formal bimetallism to a legal gold standard occurred during a paper standard (the "greenback period," of 1861–1878), with a dual legal and effective gold standard from 1879.

14.1.2.4 Periphery

The core countries attracted other countries to adopt the gold standard, in particular, British colonies and Dominions, Western European countries, and Scandinavia. These noncore countries were generally closely aligned with one or more core countries and could be viewed as constituting the "inner periphery." The "rush to the gold standard" began in the 1870s, with the adherence of Germany, France, Scandinavia, and other European countries. Legal bimetallism shifted from effective silver to effective gold monometallism around 1850, as gold discoveries in the United States and Australia resulted in overvalued gold at the mints. With silver discoveries in Nevada, the gold/silver market situation subsequently reversed itself, and to avoid a huge inflow of silver and stem an outflow of gold, many European countries suspended the coinage of silver and limited its legal-tender property. Some countries (France, Belgium, Switzerland—three founding members of the Latin Monetary Union) adopted a "limping" gold standard, in which existing former-standard silver coin retained full legal tender, permitting the monetary authority to redeem its notes in silver as well as gold.

So, while all noncore countries were in the broadly defined periphery, there is a narrower periphery: Eastern Europe, Middle East, Asia, some colonial Africa, and Latin America. These countries—including, for some purposes, also British colonies and Dominions—were in the "outer" periphery: acted on, rather than actors, in the gold standard, and generally not as committed to the gold standard. Some countries—China, Persia, parts of Latin America—never joined the classical gold standard, instead retaining their silver or bimetallic standards.

Flandreau and Jobst have a different division of noncore countries. The periphery consists of countries the currency of which has exchange-market representation only at home and possibly in one neighboring country: Dominions and colonies, Southeastern Europe, Latin America, and Asia. The periphery could also be defined as the set of countries which could not circulate abroad debt denominated in their own currency (Morys 2013, pp. 206–207).

For Flandreau and Jobst, an intermediate group consists of countries the currencies of which enjoy regional exchange-market quotations: the United States and various European countries. I find questionable the characterization of the United States as noncore. It was simply a matter of historical tradition that "the reach of the dollar-sterling exchange market extended beyond to encompass almost the entirety of American economic transactions…All the while, the balance-of-payments strength of the United States was growing, and along with it resentment of foreign-exchange dependence on London, which financed US trade even with third parties" (Officer 1996, p. 61, 63). Throughout the gold-standard heyday, the United States became more and more economically powerful, and more and more important to the international gold standard—making characterization of the country as noncore incongruous.

14.1.2.5 Why the Scramble to Gold?

The idea of a "scramble" or "rush" to gold has also been named the "monetary chain gang" (Gallarotti 1995). There was a sequential movement to gold driven by network externalities in the form of trade and investment. Dependencies, sovereign and nonsovereign, following Britain to the gold standard, are mentioned above. Similarly, German economic satellites (Netherlands, Scandinavia) followed Germany and French satellites (Switzerland, Belgium, Italy) followed France. Eichengreen and Flandreau (1996) extend this thread to India and Straits Settlements, Dutch East Indies, Korea (and, logically, Taiwan), and Philippines, following Britain, Netherlands, Japan, and United States, respectively.

The role of the fall in the price of silver in the switch to gold is subject to amendment. Perhaps it was the desire to stabilize commodity prices that was the impetus for the switch (Gallarotti 1995). And it is arguable that the fall in the price of silver relative to gold was determined by shifts in demand rather than supply (Milward 1996), whence an endogenous phenomenon. These issues warrant further attention by historians and cliometricians.

Conventional scholarly wisdom is that war indemnity helps to explain some adoptions of the gold standard (e.g., France and Germany following the Franco- Prussian War, Japan after victory over China in 1895). Industrialization is also said to play a role, with its high value/weight ratio making gold the better metal than silver for transactions large in size and volume. Also, the gold standard had a "Good Housekeeping Seal of Approval" for the inflow of long-term capital (Bordo and Rockoff 1996, of which more below). Countries with fluctuating exchange rates might have been attracted to the stability of gold. There are also political theories of gold-standard adoption. Ideologically, there was the desire to follow the monetary standard of Britain and Germany, the leading economic powers. Domestically, there was the rise of urban-capitalist and industrial over agricultural interest groups, the former favoring gold for its low inflation (Gallarotti 1995), and the perennial conflict between creditors and debtors, again the former supporting gold for its purportedly deflationary power.

Empirically unscrambling the many theories of the scramble for gold is a difficult task. Meissner (2005) adopts an approach that warrants attention and extension. Using an econometric "duration model," his determined variable is the number of years (after 1870) until a country adopts the gold standard. The strongest result is that "a country would be more likely to move to gold the more it traded with other gold standard countries" (Meissner 2005, p. 400). Also "A higher gold cover ratio [gold reserves/notes outstanding] is associated with earlier adoption times" (p. 395). So is a higher spread between domestic bond yield and British consol rate, in line with the "Good Housekeeping" hypothesis. In general, "the order in which countries adopted depended on trade patterns, financial needs, and structural constraints" (Meissner 2005, p. 401).

14.1.3 Characteristics of Gold Standard

14.1.3.1 Domestic Gold Standard

Coin Standard

Most gold-standard countries were on a coin standard (see Table 14.1). From a domestic standpoint, the coin standard had four properties. First, there was a well-defined and fixed gold content of the domestic monetary unit. For example, the dollar was defined as a specified weight of pure gold. Second, gold coin circulated as money with unlimited legal-tender power (meaning it is compulsorily acceptable means of payment of any amount in any debt transaction or obligation). Third, privately owned bullion (gold in mass, foreign coin considered as mass, or gold in the form of bars) was convertible into gold coin in unlimited amounts at the government mint or at the central bank (if one existed), and at the "mint price" (of gold, the inverse of the gold content of the monetary unit). Fourth, private parties had no restriction on their holding or use of gold (except possibly that privately created coined money could be prohibited); in particular, they may melt coin into bullion. The effect is as if coin were sold to the monetary authority (central bank, or Treasury acting as a central bank) for bullion. It sometimes made sense for the authority to sell gold bars directly for coin, even though not legally required, thus saving cost of coining. The third and fourth properties in effect committed the monetary authority to transact in coin and bullion in both directions such that the mint price, or gold content of the monetary unit approximately (because of transactions costs) governed in the marketplace.

However, even under a coin standard, gold was not the only money. Rather than a "pure" coin standard, the norm was a "mixed" coin standard, with both gold coin and other money circulating. In fact, a pure coin standard did not exist in any country during the gold-standard period. There was non-gold coin and also paper currency (notes)—issued

Table 14.2 Structure of money: major-countries aggregate

	1885	*1913*
Money supply ($ billion)	8	26
Ratio of metallic money to money supply (%)	33	15
Ratio of official reserves to money supply (%)	18	16
Ratio of official to official-plus-money gold (%)	33	54

Source Triffin (1964, p. 62)
End of year. Major countries are core (Britain, United States, France, Germany), Western Europe (Belgium, Italy, Netherlands, Sweden, Switzerland), Canada, and Japan. Money supply consists of metallic money, minor coin, paper currency, and demand deposits. Metallic money in 1885 is gold and silver coin; an overestimate, as includes commercial-bank holdings that could not be isolated from coin held outside banks by the public. Metallic money in 1913 is gold and silver coin. Official reserves are gold, silver, and foreign exchange. Official gold is gold in official reserves. Money gold is the gold-coin component of money supply

by the government, central bank, or commercial banks—and demand-deposit liabilities of banks. Generally, except for a "limping" gold standard (see above), non-gold (in particular, silver) coin was not officially convertible into gold and had only "token" status, meaning limited legal-tender power and face-value exceeding metallic value. In contrast, government or central-bank notes and central-bank deposit liabilities were directly convertible into gold coin at the fixed established price on demand. Commercial-bank notes and demand deposits might be converted not directly into gold but rather into gold-convertible government or central-bank currency. This indirect convertibility of commercial-bank liabilities would apply certainly if the government or central-bank currency were legal tender, but also generally even if it were not.

As legal tender, gold coin was always exchangeable for paper currency or deposits at the mint price, and usually the monetary authority would provide gold bars for its coin. Again, two-way transactions in unlimited amount fixed the currency price of gold at (or approximately at) the mint price. The credibility of the monetary-authority commitment to a fixed price of gold is the essence of a successful, ongoing gold-standard regime.

Over time, gold coin declined from about 1/5 of the world money supply in 1800 (2/3 for gold and silver coin together, as silver was then the predominant monetary standard) to 17% in 1885 (1/3 for gold and silver, for an eleven-major-country aggregate), and 10% in 1913 (15% for gold and silver, for the major-country aggregate) (Triffin 1964, pp. 15, 56, and see Table14.2). The main use of gold coin became not circulating medium but rather reserves for Treasuries, central banks, and (generally to a lesser extent) commercial banks.

Gold-Exchange Standard

As shown in Table 14.1, some countries in the periphery were on a gold-exchange standard, in which the monetary authority buys and sells not gold (in any form) but rather gold-convertible foreign exchange, that is, the currency of a country that itself is on the coin standard. Countries on a gold-exchange standard usually were colonies or territories of a country on a coin standard. In situations in which the periphery country lacked its own (even-coined) currency, the gold-exchange standard existed almost by default.

14.1.3.2 International Gold Standard

Properties

An "international" gold standard requires, in addition to the domestic properties, freedom both of international gold flows (private parties permitted to import or export gold without restriction) and of foreign-exchange transactions (an absence of exchange control). Then the fixed mint prices of any two countries on the gold standard imply a fixed exchange rate ("mint parity") between the countries' currencies. For example, the US mint price effective 1837 was $20.671835 (rounded) per fine ounce of gold, the British since 1717 £4.247727(+), whence dollar-sterling mint parity was $4.8665635 per pound sterling (Officer 1996, p. 51). There are actually several concepts of parity, for which (considering the dollar-sterling case) one may consult Officer (1996, Chap. 5; 2006). The lag of "legal parity" (for appraisal of British merchandise for tariffs) behind mint parity, catching up only in 1873, is an issue that warrants explanation by historians.

Gold Points and Gold-Point Arbitrage

A fixed exchange rate (at the mint parity) for two countries on the gold standard is an oversimplification, which is often made but is misleading. There were costs of importing or exporting gold. These costs included freight, insurance, handling (packing and cartage), interest on money committed to the transaction, risk premium (compensation for risk), normal profit, any deviation of purchase or sale price from the mint price, possibly mint charges, and possibly abrasion (wearing out or removal of gold content of coin—should the coin be sold abroad by weight or as bullion). Expressing the exporting costs as percent of the amount invested (or, equivalently, as percent of parity), the product of 1/100th these costs and mint parity (number of units of domestic currency per unit of foreign currency, for example, number of dollars per pound) was added to mint parity to obtain the gold-export point, the exchange rate at which gold is exported. To obtain the gold-import point, the product of 1/100th of the importing costs and mint parity was subtracted from mint parity.

If the exchange rate was greater than the gold-export point, private-sector "gold-point arbitrageurs" exported gold, thereby obtaining foreign currency. Conversely, for the exchange rate less than the gold-import point, gold was imported and foreign currency relinquished. Usually the gold was, directly or indirectly, purchased from the monetary authority of the one country and sold to the monetary authority in the other. The domestic-currency cost of the transaction "per unit of foreign currency obtained" was the gold-export point. That "per unit of foreign currency sold" was the gold-import point. Also, foreign currency was sold, or purchased, at the exchange rate. Therefore, arbitrageurs receive a profit proportional to the exchange-rate/gold-point divergence.

However, the arbitrageur supply of foreign currency eliminates profit by returning the exchange rate to below the gold-export point. Therefore, perfect "gold-point arbitrage" would ensure that the exchange rate has upper limit of the gold-export point. Similarly, the arbitrageur demand for foreign-currency returns the exchange rate to above the gold-import point, and perfect arbitrage ensures that the exchange rate has that point as a lower limit. It is important to note what induces the private sector to engage in gold-point arbitrage: (1) the profit motive and (2) the credibility of the commitment to (a) the fixed gold price and (b) freedom of foreign exchange and gold transactions, on the part of the monetary authorities of both countries.

Discussions of gold-point arbitrage are in Officer (1996, Chap. 8) and Canjels et al. (2004, pp. 871–875).

Spread, Gold-Point Estimation, and Gold-Effected Transfer of Funds

The "spread," the exchange-rate range over which arbitrage is unprofitable, is the difference between the gold-export point and gold-import point. It is sometimes convenient to express the gold points (and exchange rate) as percentage of parity. Then the spread becomes the sum of the gold points. Estimates of gold points and spreads involving

Table 14.3 Gold-point estimates

Countries	*Period*	*Gold points (%)*		*Spread (%)*	*Method of computation*
		Export	*Import*		
US/Britain	1841–1850	1.7476	3.2960	5.0436	A
US/Britain	1851–1860	1.3306	1.8631	3.1937	A
US/Britain	1881–1890	0.6585	0.7141	1.3726	A
US/Britain	1891–1900	0.6550	0.6274	1.2824	A
US/Britain	1901–1910	0.4993	0.5999	1.0992	A
US/Britain	1911–1914	0.5025	0.5915	1.0940	A
US/Britain	1879–1913	0.7706–0.1192		0.8898	B
US/Britain	1879–1913	0.4192–0.2486		0.6678	C
France/US	1877–1913	0.6888	0.6290	1.3178	D
Germany/US	1894–1913	0.4907	0.7123	1.2030	D
France/Britain	1877–1913	0.4063	0.3964	0.8027	D
Germany/Britain	1877–1913	0.3671	0.4405	0.8076	D
Germany/France	1877–1913	0.4321	0.5556	0.9877	D
Austria/Britain	1912	0.6453	0.6037	1.2490	E
Netherlands/ Britain	1912	0.5534	0.3552	0.9086	E
Scandinavia/ Britain	1912	0.3294	0.6067	0.9361	E

Sources US/Britain, 1879–1913–Canjels et al. (2004, p. 879). US/Britain, other periods–Officer (1996, p. 174). France/US, Germany/US, France/Britain, Germany/Britain, Germany/France–Morgenstern (1959, pp. 178–181). Austria/Britain, Netherlands/Britain, Scandinavia/Britain–Easton (1912, pp. 358–363)

Gold points apply to numerator country. Therefore, gold-export point is gold-import point for denominator country, and gold-import point is gold-export point for denominator country. Spread is gold-export point plus gold-import point. Scandinavia is Denmark, Sweden, and Norway

Method of computation A: sum of period-average arbitrage-cost components; B: exchange-rate behavior, nonparametric model; C: exchange-rate behavior, smooth time-trend model; D: median estimate of various authorities for various dates; E: writer's estimate. B-E: converted to percent deviation from parity; B-C: Gold-points symmetric and decline over time as shown, from beginning of period (maximum spread) to end of period (minimum spread); figure for spread is midpoint of maximum and minimum spread

core countries are presented in Table 14.3. There are many methods of obtaining or estimating gold points—Officer (1996, pp. 117–121) identifies nine techniques, which Canjels et al. (2004, p. 869) reduce to four. The main distinction is between summing cost components over time (method A in Table 14.3) and applying sophisticated time-series analysis to high-frequency, daily, exchange-rate data (methods B and C). Canjels et al. argue that their technique is superior to method A (exemplified by

Officer 1996) and that their results—especially a narrower spread than estimated by Officer—are consistent with gold-flow data. However, the Canjels et al. symmetry assumption (implying gold export and import points equidistant from parity), perhaps made for analytic convenience, is at variance with historical evidence.

Noteworthy in Table 14.3 is that the gold points, and therefore the spread, declined over time (evidenced by the dollar-sterling figures, whether methods A or B-C). Explanations involve technological improvements in transportation, communication, and arbitrage itself.

Almost always forgotten by economic historians is the fact that gold flows also were employed to transfer funds in lieu of a foreign-exchange transaction (rather than in combination with such transaction, per gold-point arbitrage). It is supremely ironic that contemporary accounts of such operations almost always pertain to gold- effected transfer of funds, whereas modern textbooks and scholarly articles deal exclusively with gold-point arbitrage! It is easy to demonstrate theoretically—and Officer shows empirically—that the spread pertinent to transfer of funds was always narrower than the gold-point arbitrage spread.

14.1.4 Implications for Money Supply and Automatic Correctives

Consider a domestic gold standard. Under a pure coin standard, gold in circulation, monetary base, and money supply are all one. With a mixed standard, the money supply is the product of the money multiplier (dependent on the commercial-banks' reserves/deposit and the nonbank-public's currency/deposit ratios) and the monetary base (the actual and potential reserves of the commercial banking system, with potential reserves held by the nonbank public). The monetary authority alters the monetary base by changing its gold holdings and its loans, discounts, and securities portfolio (non-gold assets, domestic assets). However, the level of its domestic assets is dependent on its gold reserves, because the authority generates demand liabilities (notes and deposits) by increasing its assets, and convertibility of these liabilities must be supported by a gold reserve, if the gold standard is to be maintained. Therefore, the gold standard provides a constraint on the level (or growth) of the money supply.

The international gold standard involves balance-of-payments surpluses settled by gold imports at the gold-import point and deficits financed by gold exports at the gold-export point. (Within the spread, there are no gold flows and the balance of payments is in equilibrium.) The change in the money supply is then the product of the money multiplier and the gold flow, providing the monetary authority does not change its domestic assets. For a country on a gold-exchange standard, holdings of "foreign exchange" (the reserve currency) take the place of gold. In general, the "international assets" of a monetary authority may consist of both gold and foreign exchange. Discussion of automatic correctives of a payments imbalance assumes "neutral" policy of the monetary authorities, that is, abstraction both from policies that would enhance and policies that would inhibit correction.

14.1.4.1 Traditional Mechanism

A country experiencing a balance-of-payments deficit loses gold and its money supply automatically decreases, the extent of the decrease depending on the legal or customary reserve requirements for non-gold (or non-foreign-exchange) money (for the variety of legal institutions, see Martin-Aceña 2007, p. 105). Assuming that velocity does not increase (i.e., the demand for money does not decrease), money income contracts, via the equation of exchange. Then the price level and/or real income falls. If prices are fully flexible (guaranteed only by pure *and* perfect competition), then the price level bears the full force of the deflation. As long as elasticity conditions (moderate, but typically neglected in the literature) are satisfied, exports increase not only in real but also in nominal terms, and imports similarly decrease. Symmetrically, a surplus country gains gold, the money supply increases, money income expands, the price level rises, exports decrease, and imports increase. In each case, balance-of-payments equilibrium is restored via the current account. This is called the price specie-flow mechanism; "developed in the eighteenth century, it remains the dominant approach to thinking about the gold standard today" (Eichengreen 2008, p. 24).

An extended adjustment mechanism incorporates changes in real income and interest rates. To the extent that prices are inflexible, movements of real income in the same direction as money income occur; in particular, the deficit country suffers unemployment, but the payments imbalance is nevertheless corrected.

The capital account also acts to restore balance, via the deficit-country reduced money supply increasing interest rates, inducing a net inflow of capital. The interest-rate increases also reduce real investment and thence real income and imports. Similarly, interest-rate decreases in the surplus country elicit capital outflow and increase real investment, income, and imports. This process enhances the price specie-flow current-account correction of the imbalance.

14.1.4.2 Monetary Mechanism

From a general monetarist standpoint, the traditional mechanism is unnecessary to restore payments equilibrium, because, with fixed exchange rates, gold flows simply adjust money supply to money demand. Changes in prices, real income, and interest rates are superfluous to the adjustment process. Further, under a "global-monetarist" framework, prices, interest rates, and incomes are all determined worldwide. Therefore, in logical extreme, the price-specie-flow and like mechanisms cannot even occur (on the monetarist approaches, see Kreinin and Officer 1978, Chap. 3).

For some authors (McCloskey and Richard Zecher 1976; Temin 1984, pp. 576–577; Gallarotti 1995, pp. 35–36), historical data support the monetary mechanism for the classical gold standard: Gold flows were too small to be suggestive of the traditional correctives, and prices, incomes, and interest rates moved closely in correspondence (rather than in the opposite directions predicted by the traditional adjustment mechanisms)—at least among non-outer-periphery countries, especially the core group. Hatton (1992) is skeptical of this work, while Wallace and Choudhry (1995) present evidence against global monetarism and in favor of the price specie-flow mechanism.

The "law of one price"—purchasing power parity (PPP) in weak form—is associated with the monetary approach and contravenes price specie-flow. Examining ten studies published during the period spanned by Enders (1989) and Catão and Solomou (2005), some of which are discussed in Officer (2012), I judge that eight provide at least partial support for PPP. However, PPP is generally found to be stronger over time, which leads to the open question "how long is too long for the monetary approach to receive validation?" Undoubtedly, there remains scope for additional work on automatic correctives.

14.1.5 Sources of Instability of the Gold Standard

There were three elements making for instability of the classical gold standard. First, the use of foreign exchange as reserves increased as the gold standard progressed. Available end-of-year data indicate that, worldwide, foreign exchange in official reserves (the international assets of the monetary authority) increased by 36% from 1880 to 1899 and by 356% from 1899 to 1913. In comparison, gold in official reserves increased by 160% from 1880 to 1903 but only by 88% from 1903 to 1913 (Lindert 1969, pp. 22–25). While in 1913 only Germany among the center countries held any measurable amount of foreign exchange—15% of total reserves excluding silver (which was of limited use)—the percentage for the rest of the world was double that for Germany (Table 14.4). If there were a rush to cash in foreign exchange for gold, reduction or depletion of the gold of reserve-currency countries could place the gold standard in jeopardy.

Second, Britain—the predominant reserve-currency country—was in a particularly sensitive situation. From 1899 to 1913, recorded sterling balances (mostly official) increased more than 2.5-fold (Lindert 1969, p. 22). Considering end-of-1913 data, almost half of world foreign-exchange reserves was in sterling, but the Bank of England had only 3% of world gold reserves (Tables 14.5, 14.6). Defining the "reserve ratio" of

Table 14.4 Share of foreign exchange in official reserves: 1913

	Including silver	*Excluding silver*
Britain	0	0
United States	0	0
France	0	0
Germany	13	15
Rest of world	27	31

Source Lindert (1969, pp. 10–11)
Official reserves are gold, foreign exchange, and including or excluding silver

Table 14.5 Composition of world official foreign-exchange reserves: 1913

Currency	*Percent*
British pounds	47
US dollars	2
French francs	30
German marks	16
Other	5

Source Lindert (1969, pp. 18–19)
End of year. Excludes holdings for which currency unspecified. "Other" is primarily Dutch guilders and Scandinavian kroner

Table 14.6 Official-reserves components: 1913

Country	Percent of world total	
	Gold	*Foreign exchange*
Britain	3	0
United States	27	0
France	14	0
Germany	6	5
Rest of world	50	95

Sources Gold: Board of Governors of the Federal Reserve System (1943, pp. 544–545, 551). Foreign exchange: Lindert (1969, pp. 10–11)

the reserve-currency-country monetary authority as the ratio of (i) official reserves to (ii) liabilities to foreign monetary authorities held in financial institutions in the country, in 1913 this ratio was only 31% for the Bank of England, far lower than those of the monetary authorities of the other core countries (Table 14.7).

An official run on sterling could easily force Britain off the gold standard. Because sterling was an international currency, private foreigners also held considerable liquid assets in London and could themselves initiate a run on sterling.

Third, the United States, though a core country, was a great source of instability to the gold standard. The US Treasury accumulated and held a high percentage of world gold reserves (in 1913, more than that of the three other core countries combined), resulting in an absurdly high

Table 14.7 Reserve ratio of reserve-currency countries: 1913

Country	*Including silver*	*Excluding silver*
Britain	0.31	0.31
United States	90.55	64.42
France	2.38	2.02
Germany	2.11	1.75

Source Lindert (1969, pp. 10–11, 19). Foreign-currency holdings for which currency unspecified allocated proportionately to the four currencies based on known distribution.
End-of-year ratio of official reserves to official liquid liabilities (that is, liabilities to foreign governments and central banks). Percent. Official reserves are gold, foreign exchange, and including or excluding silver

reserve ratio—Tables 14.5, 14.6, 14.7). With a decentralized banking system composed of many banks of three distinct types (national, state, savings)—De Cecco 1984, pp. 111–113) includes loan and trust companies as a fourth group—operating under different rules, but with a New York center, interbank deposits were prevalent and financial crises involving bank failures frequent. Eichengreen (1992, p. 55) sees an analogy between interior banks maintaining balances in New York and the US financial system holding sterling balances in London banks. In addition to episodic financial shocks, there was periodic, seasonal financial stress, as monies would flow back and forth between the agricultural interior and the New York banking center. Cyclically, the US demand for money shifted greatly, but the supply was relatively inelastic. This led to episodic high interest rates in the New York money market, which attracted capital from abroad. Further, there was an upward trend in the demand for money on the part of the US private sector, which exacerbated the capital inflow.

During the heyday of the gold standard, the US had no central bank to serve as a lender of last resort or otherwise help to stabilize the US monetary base. Provocatively, Officer (2002, pp. 115–117—and see Chapters 22–23) has argued that the First and Second Banks of the United States played the role of central bank, but these Banks had long vanished by the heyday of the gold standard. And the Federal Reserve had barely begun operations when the gold standard collapsed in 1914. The Treasury did not fill the void: "The US Treasury was by no means the lender of last resort of the American system; once it acquired gold, it just sat on it" (De Cecco 1984, p. 117).

Therefore, far from the United States assisting Britain, gold often flowed from the Bank of England to the United States to satisfy increases in US demand for money. Though in economic size the United States was the largest of the core countries, in many years it was a net importer rather than exporter of capital to the rest of the world—the opposite of the other core countries. The political power of silver interests (desiring to enhance the role of silver relative to gold), the accusations of farmer debtors and manufacturer exports (blaming the gold standard for deflation), and recurrent financial crises led to imperfect credibility in the US commitment to the gold standard. Runs on banks and runs on the Treasury gold reserve placed the US gold standard near collapse in the early and mid-1890s. During that period, the credibility of the Treasury's commitment to the gold standard was shaken. Indeed, the gold standard was saved in 1895 (and again in 1896) only by cooperative action of the Treasury and a bankers' syndicate, which stemmed gold exports.

Using time-series analysis on six-month commercial-bank loans (deemed a "more developed" market, because that maturity is the longest series available), Tullio and Wolters (2000, pp. 62, 67) conclude (based on previous work) that "the UK and London were more vulnerable to US influences in the period under study than they were to French and German influences" and (in their current study) that "All in all, the influence of US on UK interest rates is much stronger and lasts much longer than vice versa." The latter finding is stronger for 1897–1907 than for 1890–1896. I interpret these results as confirming the US unstable role in the gold standard.

In sum, the United States, by virtue of economic size and early experience with the gold standard, was a core country to be sure, but a core country that decidedly exacerbated the instability of the gold standard!

14.1.6 Rules of the Game

14.1.6.1 The Rules

According to the "rules of the [gold-standard] game," central banks were supposed to reinforce, rather than "sterilize" (moderate or eliminate) or ignore, the effect of gold flows on the monetary supply. A gold outflow typically decreases the international assets of the central bank and thence the monetary base and money supply. The central-bank's proper response was: (1) raise its discount rate, thereby inducing commercial banks to adopt a higher reserves/deposit ratio and therefore decreasing the money multiplier and (2) decrease lending and sell securities, thereby decreasing domestic interest-earning assets and thence the monetary base. On both counts, the money supply is further decreased. And the higher interest rate acted to increase interest rates generally and induce a capital inflow. The converse argument (involving increases in the money supply, and lower interest rates) applies symmetrically to a gold inflow.

It is interesting that the "rules of the game" did not appear in the literature until a decade after the classical gold standard ended (Eichengreen 1992, p. 36). The originator was Keynes (1925, p. 18), who wrote that given the overvalued pound upon the UK return to the gold standard and the consequent payments imbalance and incipient gold loss: "The Bank of England is *compelled* to curtail credit by all the rules of the Gold Standard game." Such "credit restriction" (money-supply decrease, in today's parlance) deflates the economy (reduces nominal GDP), reducing wages (the price level) via unemployment (decreasing real GDP). External balance is maintained at the expense of internal balance. Anticipating his *General Theory*, Keynes advocates rather an "easy credit policy" (easy-money policy) to "restore prosperity" (full employment) rather than following the rules of the game and "aggravate a depression" (worsening real GDP).

Should the central bank rather increase its domestic assets when it loses gold, it engages in "sterilization" of the gold flow and is decidedly not following the "rules of the game." The converse argument (for gold inflow) also holds, with sterilization involving the central bank decreasing its domestic assets when it gains gold.

According to the monetary approach, neither the "rules of the game" nor sterilization can have any effect except in the short run. Under fixed exchange rates, gold flows simply adjust money supply to money demand; the money supply cannot be determined by policy. The central bank can control the (reserve-asset versus domestic-asset) composition of the monetary base but not the level of the base. Indeed, the rules of the game are unnecessary in the first instance, because gold flows occur only because of a disequilibrium between money demand and money supply. When gold (or any reserve) has moved sufficiently to re-equate money supply to money demand, the gold loss or gain ceases. Thus rule (2) is unnecessary and is ineffective except possibly in the short run.

Furthermore, under global monetarism, interest rates and incomes are determined worldwide. Even core countries can influence these variables domestically only to the extent that they help determine them in the global marketplace. Therefore, rule (1) is inapplicable as well. In sum, the "rules of the game," whether followed or not, are deemed inconsequential by those who adhere to the monetary approach to the balance of payments.

14.1.6.2 Discount-Rate Rule

However, the Bank of England did, in effect, manage its discount rate ("Bank Rate") in accordance with rule (1). The Bank's primary objective was to maintain convertibility of its notes into gold, that is, to preserve the gold standard, and its principal policy tool was Bank Rate. When its liquidity ratio of gold reserves to outstanding note liabilities decreased, it would usually increase Bank Rate. The increase in Bank Rate carried with it market short-term interest rates, inducing a short-term capital inflow and thereby moving the exchange rate away from the gold-export point by increasing the exchange value of the pound. The converse would be a rise in the liquidity ratio involving a Bank Rate decrease, capital outflow, and movement of the exchange rate away from the gold import point—but if the converse held, it was in weaker form. Nevertheless, the Bank was constantly monitoring its liquidity ratio and in response altered Bank Rate almost 200 times over 1880–1913.

Time-series analyses, such as Jeanne (1995) and Davutyan and Parke (1995), essentially support that narrative. No doubt the Bank had other objectives: certainly profitability, given that it was a commercial bank (albeit with public functions), possibly, at times, economic activity ("home trade"). If Bank rate exceeded the market rate by too great a margin, the Bank's commercial business would suffer and shareholders would object. However, maintenance of note convertibility was required by law and viewed as necessary for the Bank's commercial functioning. So the goals of maintenance of convertibility and earning of satisfactory profits were not necessarily in conflict. In contrast, the studies show little concern for economic activity. The Bank's low gold holdings (which, of course, earned zero return) were viewed by contemporaries as based on an overriding concern for the interests of shareholders, that is, profitability (Gallarotti 1995, p. 115).

The Reichsbank operated in an environment similar to that of the Bank of England, except that the Reichsbank kept a greater reserve buffer and the Berlin money market was not as large as that of London. The Reichsbank, like the Bank of England, generally moved its discount rate inversely to its liquidity ratio.

However, most other central banks often violated the rule, with changes in their discount rates of inappropriate direction, or of insufficient amount or frequency. The Bank of France, in particular, kept its discount rate stable. Unlike the Bank of England, it chose to have large gold reserves (see Table 14.6), with payments imbalances accommodated by fluctuations in its gold rather than financed by short-term capital flows. This policy was due in part to a small money market in Paris. (Of course, the United States, lacking a central bank, had no discount rate to use as a policy instrument.)

14.1.6.3 Sterilization Was Dominant

As for rule (2)—that the central-bank's domestic and international assets move in the same direction—in fact the opposite behavior, sterilization,

Table 14.8 Annual changes in international and domestic assets of central bank: 1880–1913

Country	*Percent changes in same direction*
Britain	48
France	26
Germany	31
Western Europe	32
Scandinavia	40
Russia	33

Source Bloomfield (1959, p. 49)
International assets are gold, silver, and foreign exchange; domestic assets are income-earning: discounts, loans, and securities. Change in same direction implies country is following "rules of the game." Observations with zero or negligible changes in either class of assets excluded. Years when country is off gold standard excluded (see Table 14.1). Western Europe consists of Austria-Hungary, Belgium, and Netherlands; Scandinavia incorporates Denmark, Finland, Norway, and Sweden

was dominant, as shown in Table 14.8. The Bank of England followed the rule more than any other central bank, but even so violated it more often than not! The sterilization policy of the Bank of France was a substitute for discount-rate policy (Bazot et al. 2016).

14.1.6.4 Was the Bank of England Supreme?

Eichengreen quotes Keynes that the "Bank of England could almost have claimed to be the conductor of the international orchestra" (Eichengreen 1987, p. 5) and finds that Bank rate tended to lead the Reichsbank discount rate and even the Bank of France rate. Other studies confirm that the Bank of England discount rate sometimes was followed by a change in the same direction on the part of the Reichsbank, but not the reverse. And Bazot et al. (2016), with advanced time-series analysis, find that French sterilization was ultimately due to an increase in the Bank of England rate: "the Banque de France's credit to the domestic economy (discounts and advances) correlates negatively with gold flows because it correlates positively with the discount rate of the Bank of England" (p. 2).

Morys (2013), making use of central-bank archival data and sophisticated time-series analysis, examines the behavior of discount rates of 14 central banks and concludes that "a considerable amount of monetary autonomy was retained under the Classical Gold Standard, even for peripheral countries" (p. 215). However, Morys can be criticized for having no gold-flow or domestic-activity variables, for not conducting unit-root and cointegration testing, and for a principal-components solution to exchange-rate multicollinearity. Stokes and Neuburger (2016), in an ultra-sophisticated time-series analysis, determine that the London money-market rate heavily influenced the French and Reichsbank money-market rates. Their use of market rather than official interest rates brings richer data to bear on the issue of English leadership.

One concludes that the Bank of England was influential in determining foreign money-market conditions if not official rates, but Bank "leadership" or "hegemony" remains an open question and perhaps a matter of definition.

14.1.7 Stability of Gold Standard

How then did the classical gold standard cope with payments imbalances? Why was it a stable system?

14.1.7.1 Private-Sector Credibility in Convertibility

The fundamental reason for the stability of the classical gold standard is that there was always absolute private-sector credibility in the commitment to the fixed domestic-currency price of gold on the part of the center country (Britain), two (France and Germany) of the three remaining core countries, and certain other European countries (Belgium, Netherlands, Switzerland, and Scandinavia). Certainly, that was true from the late-1870s onward. For the United States, this absolute credibility applied from about 1900. In earlier periods, that commitment had a contingency aspect: it was recognized that convertibility could be suspended in the event of dire emergency (such as war); but, after normal conditions were restored, convertibility would be re-established at the pre-existing mint price and gold contracts would again be honored. The Bank Restriction Period is an example of the proper application of the contingency, as is the greenback period (even though the United States, effectively on the gold standard, was legally on bimetallism). An excellent discussion of "the gold standard as a contingent rule" is Bordo and Kydland (1996).

The absolute credibility in countries' commitment to convertibility at the existing mint price implied that there was extremely low, essentially zero, convertibility risk (the probability that Treasury or central-bank notes would not be redeemed in gold at the established mint price) and also essentially zero exchange risk (the probability that the mint parity between two currencies would be altered or that exchange control or prohibition of gold export would be instituted).

14.1.7.2 Reasons Why Commitment to Convertibility Was so Credible

There were many reasons why the commitment to convertibility was so credible.

1. Contracts were expressed in gold; if convertibility were abandoned, contracts would inevitably be violated—an undesirable outcome for the monetary authority.
2. Shocks to the domestic and world economies were infrequent and generally mild. There was basically international peace and domestic calm.
3. The London capital market was the largest, most open, most diversified in the world, and its gold market was also dominant. A high proportion of world trade was financed in sterling, London was the most important reserve-currency center, and balances of payments were often settled by transferring sterling assets rather than gold. Therefore, sterling was an international currency—not merely supplemental to gold but perhaps better: a boon to non-center countries, because sterling involved positive, not zero, interest return, and its transfer costs were much less than those of gold. Advantages to Britain were the charges for services as an international banker, differential interest returns on its financial intermediation, and the practice of countries on a sterling (gold-exchange) standard of financing payments surpluses with Britain by piling up short-term sterling assets rather than demanding Bank of England gold.
4. There was widespread ideology—and practice—of "metallist orthodoxy" and "monetary orthodoxy" (Gallarotti 1995), involving authorities' commitment to an anti-inflation, balanced-budget, stable-money policy. In particular, the ideology implied low government spending and taxes and limited monetization of government debt (financing of budget deficits by printing money). Therefore, it was not expected that a country's price level or inflation would get

out of line with that of other countries, with resulting pressure on the country's adherence to the gold standard.

5. This ideology was mirrored in, and supported by, domestic politics. Gold had won over silver, and paper and stable-money interests (bankers, industrialists, manufacturers, merchants, professionals, creditors, urban groups) over inflationary interests (farmers, landowners, miners, debtors, rural groups).
6. There was freedom from government regulation and a competitive environment, domestically and internationally. Therefore, prices and wages were more flexible than in other periods of human history (before and after). The core countries had virtually no capital controls, the center country (Britain) had adopted free trade, and the other core countries had moderate tariffs. Balance-of-payments financing and adjustment could proceed without serious impediments.
7. Internal balance (domestic macroeconomic stability, at a high level of real income and employment) was an unimportant goal of policy. Preservation of convertibility of paper currency into gold would not be superseded as the primary policy objective. While sterilization of gold flows was frequent, the purpose was more "meeting the needs of trade" (passive monetary policy) than fighting unemployment (active monetary policy).
8. The gradual establishment of mint prices over time ensured that the implied mint parities (exchange rates) were in line with relative price levels; so countries joined the gold standard with exchange rates in equilibrium.
9. Current-account and capital-account imbalances tended to be offsetting for the core countries, especially for Britain. A trade deficit induced a gold loss and a higher interest rate, reducing capital outflow and attracting a capital inflow. Indeed, the capital-exporting core countries—Britain, France, and Germany—could eliminate a gold loss simply by reducing lending abroad.

14.1.7.3 Rareness of Violations of Gold Points

Many of the above reasons not only enhanced credibility in existing mint prices and parities but also kept international-payments imbalances, and hence necessary adjustment, of small magnitude. Responding to the essentially zero convertibility and zero exchange risks implied

Table 14.9 Violations of gold points

Exchange rate	*Time period*	*Number of months*	*Violations*	
			Number	*Percent of months*
Dollar-sterling	1889–1908	240	1	0.4
Dollar-sterling	1890–1906	204	3	1.5
Franc-sterling	1889–1908	240	12	5.0
Mark-sterling	1889–1908	240	18	7.5

Sources Dollar-sterling, 1890–1906: Officer (1996, p. 235). Other: Giovannini (1993, pp. 130–131); numbers are approximate, deciphered from graph

by the credible commitment, private agents further reduced the need for balance-of-payments adjustment—via gold-point arbitrage. When the exchange rate moved beyond a gold point, arbitrage acted to return it to the spread. So it is not surprising that "violations of the gold points" were rare on a monthly average basis, as demonstrated in Table 14.9 for the dollar, franc, and mark exchange rate versus sterling. Certainly, gold-point violations did occur; but they rarely persisted sufficiently to be counted on monthly average data. Such measured violations were generally associated with financial crises.

The number of dollar-sterling violations for 1890–1906 exceeding that for 1889–1908 is due to the results emanating from different researchers using different data. Nevertheless, the important common finding is the low percent of months encompassed by violations. Canjels et al. (2004), using daily exchange-rate data, find that Officer's gold-point spread is too wide to accommodate recorded gold flows. On the other hand, Spiller and Wood (1988, p. 888), working with weekly exchange rates, conclude that "Many instances of alleged gold-point violations identified by previous authors, then, may have been nothing more than instances in which arbitrage costs may have been larger than average." This conundrum cries out for richer gold-point and gold-flow data and, of course, careful and appropriate time-series analysis.

14.1.7.4 Stabilizing Speculation

The perceived extremely low convertibility and exchange risks gave private agents profitable opportunities not only outside the spread (gold-point arbitrage) but also within the spread (exchange-rate speculation). As the exchange value of a country's currency weakened, the exchange rate approaching the gold-export point, speculators had an ever-greater incentive to purchase domestic currency with foreign currency (a short-term capital inflow); for they had good reason to believe that the exchange rate would move in the opposite direction, whereupon they would reverse their transaction at a profit. Similarly, a strengthened currency, with the exchange rate approaching the gold-import point, involved speculators selling the domestic currency for foreign currency (a short-term capital outflow). Clearly, the exchange rate would either not go beyond the gold point (via the actions of speculators of this ilk) or would quickly return to the spread (via gold-point arbitrage). Also, the further the exchange rate moved toward the gold point, the greater the potential profit opportunity, for there was a decreased distance to that gold point and an increased distance from the other point.

This "stabilizing speculation" enhanced the exchange value of depreciating currencies that were about to lose gold, and thus the gold loss could be prevented. The speculation was all the more powerful, because the absence of controls on capital movements meant private capital flows were highly responsive to exchange-rate changes. Dollar-sterling data, in Table 14.10, show that this speculation was extremely efficient in keeping the exchange rate away from the gold points—and increasingly effective over time. Interestingly, these statements hold even for the 1890s, during which at times US maintenance of currency convertibility was precarious. The average deviation of the exchange rate from the midpoint of the

Table 14.10 Average deviation of dollar-sterling exchange rate from gold-point-spread midpoint

Time period	*Percent of parity*	*Percent of spread*
Quarterly observations		
1881–1890	0.32	23
1891–1900	0.25	19
1901–1910	0.15	13
1911–1914	0.12	11
Monthly observations		
1890–1906	0.24	20

Source Officer (1996, p. 272). Year 1914 ends with second quarter spread fell decade-by-decade from about 1/3 of 1% of parity in 1881–1890 (23% of the gold-point spread) to only 12/100th of 1% of parity in 1911–1914 (11% of the spread).

Under basic target-zone theory, credibility is 100% operationally; thus, the exchange rate never violates the spread ("target zone") and exhibits "smooth pasting" at the gold points (Duarte et al. 2013). Hallwood et al. (1996) test the assumption of full credibility and find that it is a reasonable description of the sterling-franc exchange rate but that "instances of positive devaluation expectations of the dollar are…common" (Hallwood et al. 1996, p. 191), though large only during 1890–1896, consistent with the discussion in Sect. 14.1.5.

14.1.7.5 Government Policies that Enhanced Gold-Standard Stability

Government policies also enhanced gold-standard stability. First, by the turn of the century, South Africa—the main world gold producer—sold all its gold in London, either to private parties or actively to the Bank of England, with the Bank serving also as residual purchaser of the gold. Thus, the Bank had the means to replenish its gold reserves. Second, the orthodox-metallism ideology and the leadership of the Bank of England—other central banks would often gear their monetary policy to that of the Bank—kept monetary policies harmonized. Monetary discipline was maintained.

Third, countries used "gold devices," primarily the manipulation of gold points, to affect gold flows. Consider the Bank of England. By law, the Bank had to redeem its notes in domestic gold coin (sovereigns) at a minimum price equivalent to £3 17 s. 10½d. per standard ounce of gold and purchase gold bars at a minimum price of £3 17 s. 9d. Beyond that, the Bank had tremendous discretion. It would foster gold imports by lowering the foreign gold-export point (British gold-import point, number of units of foreign currency per pound) through interest-free loans to gold importers or raising its purchase price for bars and foreign coin. The Bank would discourage gold exports by lowering the foreign gold-import point (British gold-export point) via increasing its selling prices for gold bars and foreign coin, refusing to sell bars, or redeeming its notes in underweight domestic gold coin. These policies were alternative to increasing Bank Rate.

The US Treasury followed similar policies at times. In addition to providing interest-free loans to gold importers and changing the premium at which it would sell bars (or refusing to sell bars outright), the Treasury condoned banking syndicates to put pressure on gold arbitrageurs to desist from gold export in 1895 and 1896, a time when the US adherence to the gold standard was under stress. Officer (1996, Chap. 9) provides detailed data on the two countries' gold devices from a bilateral standpoint.

The Bank of France and Reichsbank employed gold devices relative to discount-rate changes more than Britain did. Some additional policies included converting notes into gold only in Paris or Berlin rather than at branches elsewhere in the country, the Bank of France converting its notes in silver coin rather than gold (permitted under its "limping" gold standard), and the Reichsbank using moral suasion to discourage the export of gold. Gold devices combined with a huge buffer stock of gold enabled the Bank of France to keep its discount rate stable while maintaining convertibility. In the 1900s, the Bank stopped the use of gold devices, replacing them with foreign-exchange market intervention (Bazot et al. 2016).

Also, the monetary system was adept at conserving gold, as evidenced in Table 14.2. This was important, because the increased gold required for a growing world economy could be obtained only from mining or from non-monetary hoards. While the money supply for the eleven-major-country aggregate more than tripled from 1885 to 1913, the percent of the money supply in the form of metallic money (gold and silver) more than halved. This process did not make the gold standard unstable, because gold moved into commercial-bank and central-bank (or Treasury) reserves: the ratio of gold in official reserves to official plus money gold increased from 33 to 54%. The relative influence of the public versus private sector in reducing the proportion of metallic money in the money supply is an issue warranting exploration by monetary historians.

Further, while the stable environment in which the gold standard operated did not require regular central-bank cooperation, such cooperation was forthcoming when needed, that is, during financial crises. Although Britain was the center country, the precarious liquidity position of the Bank of England meant that it was more often the recipient than the provider of financial assistance. In crises, it would obtain loans from the Bank of France (also on occasion from other central banks), and the Bank of France would sometimes purchase sterling to push up that currency's exchange value. "Interestingly, it was because France cared so much more about domestic finance, that it came to care about international finance…International markets represented the first line of defense for French finance; i.e. mitigating the problem at the source" (Eichengreen 2008, p. 33). Assistance also went from the Bank of England to other central banks, as needed. And cooperation went beyond the core countries. "In effect, the resources on which any one country could draw when its gold parity was under attack extended beyond its own reserves to those that could be borrowed from other gold-standard countries" (Eichengreen 2008, p. 33). Further, the credible commitment was so strong that private bankers did not hesitate to make loans to central banks in difficulty. Cooperation during the gold standard is discussed by Gallarotti (1995, Chap. 3) and Eichengreen (1992, pp. 48–52).

In sum, "virtuous" two-way interactions were responsible for the stability of the gold standard. The credible commitment to convertibility of paper money at the established mint price, and therefore the fixed mint parities, were both a cause and a result of (1) the stable environment in which the gold standard operated, (2) the stabilizing behavior of arbitrageurs and speculators, and (3) the responsible policies of the authorities—and (1), (2), and (3), and their individual elements, also interacted positively among themselves.

14.1.8 Experience of Periphery

An important reason for periphery countries to join and maintain the gold standard was the access to the capital markets of the core countries thereby fostered. Adherence to the gold standard connoted that the peripheral country would follow responsible monetary, fiscal, and debt-management policies—and, in particular, faithfully repay the interest on and principal of debt. This "Good Housekeeping Seal of Approval," by reducing the risk premium, involved a lower interest rate on the country's bonds sold abroad, and very likely a higher volume of borrowing. The favorable terms and greater borrowing enhanced the country's economic development. However, Flandreau and Zumer (2004) argue and demonstrate that gold-standard adherence was unimportant in explaining international interest-rate spreads. Rather, a country's debt burden (ratio of interest-service to revenue) and default history were the crucial explanatory variables. This finding detracts from the alleged advantage of the gold standard to the periphery.

Furthermore, periphery countries bore the brunt of the "burden of adjustment" of payments imbalances with the core (and other Western

European) countries, in four ways. First, some of the periphery countries were on a gold-exchange standard. When they ran a surplus, they typically increased (and with a deficit, decreased) their liquid balances in London (or other reserve-currency country) rather than withdraw gold from (or ship gold to) the reserve-currency country. The monetary base of the periphery country would increase, or decrease, but that of the reserve-currency country would remain unchanged. This meant that such changes in domestic variables—prices, incomes, interest rates, portfolios, etc.—that occurred to correct the surplus or deficit were primarily in the periphery country. The periphery, rather than the core, bore the burden of adjustment.

Second, the non-gold (silver and inconvertible-paper) periphery was subject to substantial exchange-rate variability, which altered real exchange rates, generating core-periphery payments adjustment. "From the perspective of the core, exchange rate flexibility in the periphery facilitated international relative price adjustment, while maintaining the monetary stability required for the preservation of the gold peg" (Catão and Solomou 2005, p. 1272).

Third, when Bank Rate increased, London drew funds from France and Germany, which attracted funds from other Western European and Scandinavian countries, which drew capital from the periphery. Also, it was easy for a core country to correct a deficit by reducing lending to, or bringing capital home from, the periphery, thus bringing about "sudden stops" to the capital inflow of periphery countries.

Fourth, the periphery countries were underdeveloped; their exports were largely primary products (agriculture and mining), which inherently were extremely sensitive to world market conditions. This feature made adjustment in the periphery compared to the core take the form more of real than financial correction. This conclusion also follows from the fact that capital obtained from core countries for the purpose of economic development was subject to interruption and even reversal ("sudden stops"). While the periphery was probably better off with access to the capital than in isolation, its welfare gain was reduced by the instability of capital import.

Fifth, peripheral countries were subjected to financial crises more than the core. Bordo and Meissner (2011, p. 85) show that "higher capital inflows were strongly related to a higher probability of having any kind of crisis." They examine the roles of "original sin" (hard-currency-denomination debt), "currency mismatches" (lack of assets-liabilities offsets in foreign-currency debt), and "debt intolerance" (past defaults) in 30 countries over 1880–1913. Their results "tend to confirm that it is difficult to find robust determinants of financial crises." However, they find a strange quadratic relationship between "the ratio of hard-currency debt to total debt" and debt crises. Economic historians look for patterns, of which either the absence or the weirdness is frustrating: the answer generally lies in further research.

The experience on adherence to the gold standard differed among periphery groups. The important British Dominions and colonies—Australia, New Zealand, Canada, and India—successfully maintained the gold standard. They were politically stable and, of course, heavily influenced by Britain. They paid the price of serving as an economic cushion to the Bank of England's financial situation; but, compared to the rest of the periphery, gained a relatively stable long-term capital inflow. The European periphery had the advantage of emigrant remittances, which, according to Esteves and Khoudour-Castéras (2009, p. 980), served as a substitute for capital inflows and "were instrumental in relieving [international] credit constraints to developing nations." Some European periphery countries "shadowed the gold standard." Even with inconvertible paper currency, they maintained relatively stable exchange rates and prices, thus largely behaving as if they were on the gold standard (see, for example, Tattara and Volpe 1997, Tattara 2003, Martin-Aceña et al. 2012).

In undeveloped Latin American and Asia, adherence to the gold standard was fragile, with lack of complete credibility in the commitment to convertibility. Many of the reasons for credible commitment that applied to the core countries were absent—for example, there were powerful inflationary interests, strong balance-of-payments shocks, and rudimentary banking sectors. For Latin America and Asia, the cost of adhering to the gold standard was very apparent: loss of the ability to depreciate the currency to counter reductions in exports. Yet the gain, in terms of a steady capital inflow from the core countries, was not as stable or reliable as for the British Dominions and colonies.

Comparisons of periphery-country experience with the core and with each other are presented in Table 14.11. It is perhaps surprising that Southern Europe exhibits even less adherence to the gold standard than does Latin America. In terms of money growth, there is a schism (seen most clearly in the coefficient of variation—ratio of standard deviation to mean) between stability for the core, Scandinavia, Western Europe, and

Table 14.11 Country-group statistics: 1881–1913

Country group	*Component-Country Means*						
	Gold-Standard adherence	*Money growth*		*Government deficit*		*Inflation*	
	Mean	*Mean*	*St. Dev.*	*Mean*	*St. Dev.*	*Mean*	*St. Dev.*
Core	100	4.0	3.8	0.8	0.9	0.3	3.4
Scandinavia	100	5.6	4.2	0.3	0.7	0.4	3.2
Western Europe	100	4.2	3.5	2.1	0.3	0.6	3.6
Dominions	100	5.4	5.5	7.8	1.6	0.4	2.6
Southern Europe	22	2.5	6.7	0.6	0.7	0.3	2.6
Latin America	32	3.3	15.9	2.3	3.1	4.0	13.6
Japan	52	7.2	14.5	−3.1	3.3	4.6	5.5

Sources Gold-Standard adherence: Table 1. Other columns: Bordo and Schwartz (1996, pp. 46–47, 52–53, 58–59)

Gold-standard adherence is percent of years 1881–1913 on gold standard. Money growth is the time coefficient from annual regression of natural logarithm of M2 on constant and time trend. Government deficit is percent of GNP. Inflation is the time coefficient from annual regression of natural logarithm of GDP deflator (or equivalent) on constant and time trend

Core: Britain, United States, France, Germany. Scandinavia: Denmark, Finland, Norway, Sweden. Western Europe: Belgium, Netherlands, Switzerland. Dominions: Australia, Canada. Southern Europe: Italy, Portugal, Spain. Latin America: Argentina, Brazil, Chile

Dominions, versus instability for the remaining periphery. The figures for government deficit have some anomalies; but the core, Scandinavia, and Western Europe certainly exhibit “monetary orthodoxy.” Inflation level and variability are relatively high for Latin America and Japan. The figures do not uniformly reflect the quantity theory of money; but, except for the Southern Europe anomaly, there is broadly an association between gold-standard adherence and stable money in all senses.

14.1.9 Performance

Performance of the gold standard is reasonably evaluated via contrast with alternative international monetary systems, whether past or future—and the possible criteria are various. Consider first, in Table 14.12, monetary criteria for the US heyday gold standard (1879–1913) in comparison with previous US systems: First and Second Banks (1792–1810, 1817–1838),

Table 14.12 US monetary statistics: gold standard versus other periods

Period	*Exchange-market pressure (period mean, percent)*		*Ratio of monetary base to specie stock (end of year)*	
	Algebraic value	*Absolute value*	*Mean*	*Coeff. of variation (%)*
1792–1810	0.83	7.09	1.22	6.86
1811–1816	7.04	8.20	1.25	17.80
1817–1838	2.74	6.51	1.27	11.45
1839–1846	4.89	7.76	1.06	4.69
1847–1861	−17.00	17.89	1.00	10.81
1862–1878	−10.26	10.27	3.72	42.77
1879–1913	0.63	2.69	2.17	14.80

Source Officer (2002, p. 135, 137)
Statistics of annual values

Independent Treasury (1847–1861), intermittent paper standards (1811–1816, 1839–1846), and greenback period (1862–1878). Very revealing, but almost neglected in the historical literature, is exchange-market pressure (EMP) as a criterion of performance. Under certain assumptions (no money illusion, money-market equilibrium, purely monetary model, small open economy—assumptions Officer 2002, p. 134, defends for the US gold-standard period), EMP in favor of the domestic currency (US dollar) is the unweighted sum of "payments imbalance as percentage of the monetary base" and "percentage change in the foreign-currency price of the dollar." Whether taking the mean of algebraic or absolute values of EMP as the criterion, EMP is lowest for the gold standard—in fact, by a multiple in 11 of 12 comparisons. At least for the United States and for whatever reasons, the classical gold standard worked to minimize exchange-market pressure better than all previous alternatives.

The other monetary criterion is decidedly unfavorable to the gold standard. The ratio of the monetary base to specie stock ("pyramiding ratio") measures discipline in restricting the monetary base. Under a pure coin standard, the ratio is unity; so the ideal ratio is a zero coefficient of variation around a unitary mean. It is to be expected that the greenback period is least disciplined, but the gold standard follows as second (mean) or third (coefficient of variation) highest pyramiding ratio. Paradoxically, the flexible ratio may help to explain the high gold-standard EMP efficiency. It might also reflect the unstable role of the United States in the working of the gold standard!

Restricting comparisons to later and different monetary systems, Tables 14.13 and 14.14 measure inflation and real per-capita income growth for the four core countries. In mean inflation, the gold-standard is tops, with all four countries having lower inflation than under Bretton Woods or floating exchange rates. However, for no country is the variability of inflation lowest. And for no country does the gold standard entail maximum mean growth—Bretton Woods exhibits highest mean

Table 14.13 Inflation in core countries: gold standard versus later periods

Country	*Gold standard (1881–1913)*		*Bretton Woods (1946–1970)*		*Floating Exchange rates (1974–1995)*	
	Mean	*Standard Deviation*	*Mean*	*St. Dev.*	*Mean*	*St. Dev.*
Britain	0.3	3.1	3.9	2.2	7.5	5.6
United States	0.3	3.0	2.5	3.5	5.0	2.4
France	0.0	4.9	5.0	3.5	6.4	3.8
Germany	0.6	2.6	2.7	4.0	3.2	1.3

Source Bordo and Schwartz (1999, p. 205)
Mean inflation is time coefficient from annual regression of natural logarithm of GDP deflator (or equivalent) on constant and time trend. For United States, gold-standard mean is 0.4 using alternative data

Table 14.14 Growth of core countries: gold standard versus later periods

Country	*Gold standard (1881–1913)*		*Bretton Woods (1946–1970)*		*Floating Exchange Rates (1974–1995)*	
	Mean	*Standard Deviation*	*Mean*	*St. Dev.*	*Mean*	*St. Dev.*
Britain	1.1	2.4	2.1	1.8	1.8	2.3
United States	1.8	4.9	2.0	4.6	1.5	2.3
France	1.5	4.7	4.1	2.1	1.7	1.5
Germany	1.7	2.8	5.0	3.3	1.1	4.9

Source Bordo and Schwartz (1999, p. 205)
Mean growth is time coefficient from annual regression of natural logarithm of real per-capita GDP (or equivalent) on constant and time trend. For United States, gold-standard mean and standard deviation are 1.6 and 2.7, respectively, using alternative data

growth—or minimum standard deviation of growth. Of course, historical time periods reflect more than differential monetary systems; but the gold-standard balancing of relatively low inflation with relatively low growth is suggestive of a trade-off offered by that system.

Representative of pertinent time-series-analysis literature is the careful study for the gold-standard core provided by Bordo et al. (2010), who present a somewhat different conclusion. They distinguish "good deflation" (induced by positive supply shocks) from "bad deflation" (resulting from negative demand shocks). Although they do not make comparisons with other periods in this paper, they do see supply shocks (productivity improvements) as having significant effects on growth. Their time-series analysis shows a structural break around 1896, whereby deflation preceded inflation. For the European core, money is essentially neutral; but for the United States, monetary shocks significantly affect output—not unexpected for this core country.

14.1.10 Breakdown of Gold Standard

The classical gold standard was at its height at the end of 1913, ironically just before it came to an end. The proximate cause of the breakdown of the classical gold standard was political: the advent of World War I in August 1914. However, it was the Bank of England's precarious liquidity position and the gold-exchange standard that were the underlying cause. With the outbreak of war, a run on sterling led Britain to impose extreme exchange control—a postponement of both domestic and international payments—that made the international gold standard non-operational. Convertibility was not legally suspended; but moral suasion, legalistic action, and regulation had the same effect. Gold exports were restricted by extralegal means (and by Trading with the Enemy legislation), with the Bank of England commandeering all gold imports and applying moral suasion to bankers and bullion brokers.

Almost all other gold-standard countries undertook similar policies in 1914 and 1915. The United States entered the war and ended its gold standard late, adopting extralegal restrictions on convertibility in 1917 (although in 1914 New York banks had temporarily imposed an informal embargo on gold exports). An effect of the universal removal of currency convertibility was the ineffectiveness of mint parities and inapplicability of gold points: floating exchange rates resulted.

The classical gold standard possessed strong elements both of stability and instability. In the end, the shock of war led to dominance of the unstable forces. It is an open question how long the gold standard would have lasted had World War I not brought it to a close, that is, whether and when the forces making for instability would have overcome those supporting stability.

References

Bazot G, Bordo MD, Monnet E (2016) International shocks and the balance sheet of the Bank of France under the classical gold standard. Explor Econ Hist 62:87–107.

Bloomfield AI (1959) Monetary policy under the international gold standard: 1880–1914. Federal Reserve Bank of New York, New York.

Board of Governors of the Federal Reserve System (1943) Banking and monetary statistics 1914–1941.

Bordo MD (1993) The gold standard, Bretton Woods and other monetary regimes: A historical appraisal. Fed Reserve Bank St Louis Rev 75(2):123–191.

Bordo MD, Kydland FE (1996) The gold standard as a commitment mechanism. In: Bayoumi T, Eichengreen B, Taylor MP (eds) Modern perspectives on the gold standard. Cambridge University Press, Cambridge, pp 55–100.

Bordo MD, Meissner CM (2011) Foreign capital, financial crises and incomes in the first era of globalization. Eur Rev Econ Hist 15(1):61–91.

Bordo MD, Rockoff H (1996) The gold standard as a 'good housekeeping seal of approval.' J Econ Hist 56(2):389–428.

Bordo MD, Schwartz AJ (1996) The operation of the specie standard. In: de Macedo JB, Eichengreen B, Reis J (eds) Currency convertibility. Routledge, London, pp 11–83.

Bordo MD, Schwartz AJ (1999) Monetary policy regimes and economic performance: The historical record. In: Taylor JB, Woodford M (eds) Handbook of Macroeconomics, vol 1A. Elsevier, Amsterdam, pp 149–234.

Bordo MD, Landon-Lane J, Redish A (2010) Deflation, productivity shocks and gold: Evidence from the 1880-1914 period. Open Econ Rev 21(4):515–546.

Canjels E, Prakash-Canjels G, Taylor AM (2004) Measuring market integration: Foreign exchange arbitrage and the gold standard, 1879–1913. Rev Econ Stat 86(4):868–882.

Catão LAV, Solomou SN (2005) Effective exchange rates and the classical gold standard adjustment. Am Econ Rev 95(4):1259–1275.

Conant CA (1915) A history of modern banks of issue. Putnam's, New York.

Davutyan N, Parke WR (1995) The operations of the Bank of England, 1890-1908: A dynamic probit approach. J Money Credit Bank 27(4):1099–1112.
De Cecco M (1984) The international gold standard. St. Martin's, New York.
Dimitrova K, Fantacci L (2010) The establishment of the gold standard in South-east Europe: Convergence to a new system or divergence from an old one? In: Baubeau P, Ögren A (eds) Convergence and divergence of National Financial Systems. Pickering & Chatto, London, pp 179–196.
Duarte AP, Andrade JS, Duarte A (2013) Exchange rate target zones: A survey of the literature. J Econ Surv 27(2):247–268.
Easton HT (1912) Tate's modern cambist. Effingham Wilson, London.
Eichengreen B (1987) Conducting the international orchestra: Bank of England leadership under the classical gold standard. J Int Money Financ 6(1):5–29.
Eichengreen B (1992) Golden fetters. Oxford University Press, Oxford.
Eichengreen B (2008) Globalizing capital, 2nd edn. Princeton University Press, Princeton.
Eichengreen B, Flandreau M (1996) The geography of the gold standard. In: de Macedo JB, Eichengreen B, Reis J (eds) Currency convertibility. Routledge, London, pp 113–143.
Enders W (1989) Unit roots and the real exchange rate before World War I: The case of Britain and the USA. J Int Money Finance 8(1):59–73.
Esteves R, Khoudour-Castéras D (2009) A fantastic rain of gold: European migrants' remittances and balance of payments adjustment during the gold standard period. J Econ Hist 69(4):951–985.
Flandreau M, Jobst C (2005) The ties that divide: a network analysis of the international monetary system, 1890–1910. J Econ Hist 65(4):977–1007.
Flandreau M, Le Cacheux J, Zumer F, Dornbusch R, Honohan P (1998) Stability without a pact? Lessons from the European gold standard, 1880–1914. Econ Policy 13(26):115–162.
Flandreau M, Zumer F (2004) The making of global finance 1880–1913. OECD, Paris.
Gallarotti GM (1995) The anatomy of an international monetary regime. Oxford University Press, Oxford.
Giovannini A (1993) Bretton woods and its precursors: Rules versus discretion in the history of international monetary regimes. In: Bordo MD, Eichengreen B (eds) A retrospective on the Bretton Woods system. University of Chicago Press, Chicago, pp 109–147.
Hallwood CP, MacDonald R, Marsh IW (1996) Credibility and fundamentals: Were the classical and interwar gold standards well-behaved target zones? In: Bayoumi T, Eichengreen B, Taylor MP (eds) Modern perspectives on the gold standard. Cambridge University Press, Cambridge, pp 129–161.

Hatton TJ (1992) Price determination under the gold standard: Britain 1880-1913. In: Broadberry SN, Crafts NFR (eds) Britain in the international economy. Cambridge University Press, Cambridge, pp 137–156.
Jeanne O (1995) Monetary policy in England 1893-1914: A structural VAR analysis. Explor Econ Hist 32(3):302–326.
Keynes JM (1925) The economic consequences of Mr. Churchill. Hogarth, London.
Kreinin ME, Officer LH (1978) The monetary approach to the balance of payments: A survey. International Finance Section, Princeton University, Princeton.
Lindert PH (1969) Key currencies and gold, 1900–1913. International Finance Section, Princeton University, Princeton.
Martin-Aceña P (2007) The gold standard: A review from the periphery. In: Cottrell PL, Notaras G, Tortella G (eds) From the Athenian tetradrachm to the euro. Ashgate, Aldershot, pp 96–112.
Martin-Aceña P, Ruiz EM, Nogues-Marco P (2012) Floating against the tide: Spanish monetary policy, 1870-1931. In: Ögren A, Oksendal LF (eds) The gold standard peripheries. Palgrave Macmillan, London, pp 145–173.
McCloskey DN, Richard Zecher J (1976) How the gold standard worked, 1880-1913. In: Frenkel JA, Johnson HG (eds) The monetary approach to the balance of payments. University of Toronto Press, Toronto, pp 357–385.
Meissner CM (2005) A new world order: Explaining the international diffusion of the gold standard, 1870-1913. J Int Econ 66(2):385–406.
Milward AS (1996) The origins of the gold standard. In: de Macedo JB, Eichengreen B, Reis J (eds) Currency convertibility. Routledge, London, pp 87–101.
Mitchener KJ, Weidenmier MD (2015) Was the classical gold standard credible on the periphery? Evidence from currency risk. J Econ Hist 75(2):479–511.
Morgenstern O (1959) International financial transactions and business cycles. Princeton University Press, Princeton.
Morys M (2013) Discount rate policy under the classical gold standard: Core versus periphery (1870s-1914). Explor Econ Hist 50(2):205–226.
Officer LH (1996) Between the Dollar-Sterling gold points. Cambridge University Press, Cambridge.
Officer LH (2002) The U.S. specie standard, 1792–1932: Some monetarist arithmetic. Explor Econ Hist 39(2):113–153.
Officer LH (2006) Dollar-sterling parity: 1789–1978. In Carter SB et al (eds) Historical statistics of the United States, millennial edn. pp 5–561-563.
Officer LH (2008) Gold standard. In: Whaples R (ed) EH.Net encyclopedia.
Officer LH (2012) Purchasing power parity in economic history. In: James J, Marsh IW, Sarno L (eds) Handbook of exchange rates. Wiley, Hoboken, pp 161–187.

Spiller PT, Wood RO (1988) Arbitrage during the Dollar-Sterling gold standard, 1899–1908: An econometric approach. J Polit Econ 96(4):882–892.
Stokes HH, Neuburger HM (2016) The “thin film of gold”—Interest rates and gold flows in the classical gold-standard era (unpublished).
Tattara G (2003) Paper money but a gold debt: Italy in the gold standard. Explor Econ Hist 40(2):122–142
Tattara G, Volpe M (1997) Italy, the fiscal-dominance model, and the gold-standard age. In: Marcuzzo MC, Officer LH, Rosselli A (eds) Monetary standards and exchange rates. Routledge, London, pp 229–263.
Temin P (1984) Comment. In: Bordo MD, Schwartz AJ (eds) A retrospective on the classical gold standard, 1821–1931. University of Chicago Press, Chicago, pp 576–581.
Triffin R (1964) The evolution of the international monetary system. International Finance Section, Princeton University, Princeton.
Tullio G, Wolters J (2000) Interest rate linkages between the US and the UK during the classical gold standard. Scottish J Polit Econ 47(1):61–71.
Wallace MS, Choudhry T (1995) The gold standard: Perfectly integrated world markets or slow adjustment of prices and interest rates? J Int Money Financ 14(3):349–371.

CHAPTER 15

Bretton Woods System

15.1 A Retrospective on the Bretton Woods System: Lessons for International Monetary Reform

Originally published in *Journal of Economic History*, Vol. 53, No. 3 (September 1993): pp. 672–673, copyright Cambridge University Press. Reprinted with permission.

A Retrospective on the Bretton Woods System: Lessons for International Monetary Reform. Edited by Michael D. Bordo and Barry Eichengreen. Chicago: University of Chicago Press, 1993. pp. xiii, 675.

What can one say about a 26-year-old international monetary system that died? That its name and dates resound (Bretton Woods [BW], 1946–1971). That the National Bureau for Economic Research held a conference right at the hotel in Bretton Woods, New Hampshire, where the negotiators met in 1944 (Mt. Washington Hotel). That the conference proceedings be preserved in a massive volume (675 pages). That the editors of the volume be two outstanding international monetary historians (Michael Bordo and Barry Eichengreen). That a magnificent introductory chapter be written by the first editor ("Michael D. Bordo's comprehensive assessment of the BW system leaves no stone unturned—history from A to Z, performance measures, hopes at the beginning and

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_15

failure at the end," says Rudiger Dornbusch [p. 99].) That an equally impressive concluding chapter be provided by the second editor (Eichengreen asks all the relevant questions, considers the answers given by the conference participants, offers his own views, and, as always, offers some new historical insights and suggestive econometric results).

The organizers of the conference deserve credit for assembling a magnificent array of scholars. A nice feature of the conference and the resulting book is that the authors of papers tend to be "younger," whereas the commentators and panel discussants are of an "older generation." The latter are prone to reminisce, but they provide useful information to the historian. In particular, Edward Bernstein, who was there, discusses how the International Monetary Fund (IMF) viewed postwar payments problems and also the attitude of the Soviet Union at Bretton Woods. He notes the "costly detour" traveled by the Soviet Union in rejecting the IMF at that time and (in the form of component parts of the country's break-up) later seeking to join (then joining) the organization almost half a century later.

An issue for readers of this JOURNAL is whether the BW regime is "history" or just a "recent event." Eichengreen wisely comments that "Historians following too close on the heels of events, it is said, risk getting kicked in the teeth" (p. 621). Still, there has been time for a certain perspective. This enables Bordo to set the tone by dividing the BW era into two subperiods: 1946–1958 (preconvertible, before Western Europe embraced current-account convertibility) and 1959–1970 (convertible). What is or is not "history" may be generation specific. Those of us who were in graduate school during the BW period view the interwar period with the same historical sense as younger observers may view BW.

Is the volume an obituary for BW? Opinions differ among the conference participants. Rudiger Dornbusch and Richard Cooper argue that BW continues to live. Only in the sense of fixed parities and a gold base is the system dead. The BW institutional structure, many of its rules, and its accomplishment of the liberalization of trade and payments survive. Robert Mundell takes the view that the BW system never lived: "The Bretton Woods Agreement accommodated the rest of the world to an international monetary system that already existed" (p. 605)—namely, the gold-dollar standard determined by the U.S.-U.K.-France Tripartite Agreement of 1936 and reinforced by World War II and the resulting inconvertibility of European currencies.

Martin Feldstein calls BW "a system that never was" (p. 613). As Ronald McKinnon states, there was a dichotomy between the IMF Articles of Agreement negotiated at Bretton Woods and the so-called BW system. The Articles envisaged a symmetrical exchange-rate system; further, it supported national macroeconomic autonomy, permitting countries to choose their output-inflation point unconstrained by internationally set rules. Instead, an asymmetrical dollar standard evolved. Bordo also takes this point of view: "The system that began operations after the Bretton Woods conference and the establishment of the IMF was different in many major respects from what the architects intended…Instead of a community of equal currencies managed by the IMF, the system was managed by the United States in cooperation with the other members of the G10" (pp. 37, 74).

The volume contains a variety of contributions. There are nice theoretical studies by Alan Stockman, analyzing the international transmission of disturbances, and Kathryn Dominguez, who offers a game-theoretic approach to the role of the IMF. Sebastian Edwards (with Julio Santaella) presents another of his many—and worthwhile—papers on devaluation experiences of developing countries. G. John Ikenberry takes a political-science viewpoint to BW, with less emphasis on the Keynes versus White Plan conflict and more attention to the "community of experts composed of liberally minded British and American economists and policy specialists who shared a set of technical and normative views about the desirable features of the international monetary order and who were given remarkable autonomy to negotiate a deal" (p. 177).

For historians, the chapters that examine BW empirically relative to the classical gold standard and other periods are of greatest interest. Bordo again sets the stage by performing macroeconomic comparisons over a century, distinguishing the classical gold standard (1881 to 1913) from the interwar (1919 to 1938), pre-convertible BW (1946 to 1958), convertible BW (1959 to 1970), and post-BW floating exchange rate (1974 to 1989) periods. "The [convertible] Bretton Woods regime exhibited the best overall macro performance of any regime" (p. 27). Various explanations are possible; but Eichengreen demonstrates that, at least for the United Kingdom, aggregate-demand and aggregate-supply disturbances were milder than in the other periods.

Very interesting is Alberto Giovannini's derivation of limits on the domestic interest rate set by credible gold points, foreign interest rate, and spot exchange rate. Violations are rare in the classical gold standard, suggesting that the gold points were credible. Less impressive is Giovannini's direct testing for gold-point violations. He repeats mistakes made by Truman Clark ("Violations of the Gold Points, 1890–1908," *Journal of Political Economy*, 92, Oct. 1984) in assuming that prewar gold-point arbitrageurs used cable transfers and took out forward contracts to cover exchange-rate risk (whereas in fact demand bills were their predominant exchange-rate instrument); that the direct costs of arbitrage were constant for long periods (whereas in fact they changed at least month to month); and that neglect of commissions in estimated arbitrage cost explains gold-point violations (whereas in fact banks avoided commission expense by using agents abroad). Giovannini's work is complemented by Richard Marston's study of interest differentials during BW and the post-BW float.

The conference participants devote much attention to the reasons for the collapse of BW. The broad consensus (Bordo, Alexander Swoboda, W. Max Corden, probably Maurice Obstfeld) is that when the dollar became substantially overvalued and Europe/Japan refused to adjust, the United States forced the issue by refusing to deflate (or disinflate) and sought to achieve a real devaluation via a nominal one—though a minority opinion (Peter Garber) focuses on the confidence rather than the adjustment problem.

Have the lessons of BW been learned? Those participants that mention the European Monetary System (EMS) look on it as a successful experience. Yet is not the EMS analogous to BW—adjustable pegs, a hegemonic country (Germany playing the role of the United States), and external constraints on domestic macroeconomic policy? Just as Europe/Japan would not subordinate their policies to U.S.-relative inflation, European countries do not wish to adjust to German-relative disinflation. What can one say about a 26-year-old international monetary system that died? Let it rest in peace.

CHAPTER 16

Afterword to Part IV

16.1 Metallic Standards (Chapter 13)

The essays on metallic standards—Gold Standard, Silver Standard, Bimetallism—originally appeared in the second edition of *The New Palgrave*.[1] It is instructive to see how they relate to comparable entries in the original Palgrave.[2] These entries are Gold as Standard, Gold Bullion as a Commodity at the Mints, Gold Points in Foreign Exchanges, Silver as Standard, Silver Legislation in the United States, Bi-Metallism, Latin Union. Excellent bibliographies in the entries reveal the contemporary state of the art.

Commonality with the present essays is threefold. First, there are neat outlines of institutional conditions required for a gold-coin standard and for bimetallism, with a brief account for a silver standard. Mint regulations that underlie the gold standard in England, France, Germany, Italy, and the United States are presented. Note that these are the four "core countries" plus Italy. There is also a comprehensive description of the treaty regulations of the Latin Monetary Union regarding gold and silver money. Discussion of U.S. silver legislation is purely institutional.

Second, there is recognition of instability of each of the three standards. Mention is made of the evolution of a pure gold-coin standard to a mixed standard in which "gold, the international standard of value, still circulates in considerable quantities, but it chiefly serves as a reserve to support the fiduciary currency and the book-entry transactions which to

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_16

a large extent replace it in ordinary use" (Palgrave 1896, p. 222). Instability (cited as "disadvantages") can result from a sudden increase in the demand for gold, resulting in non-fulfilment of promises to pay. The word "instability" is used in connection with long-run changes in the supply of gold, with resulting changes in the "value of gold" [inverse of the price level], resulting in debtor-versus-creditor gains or losses.

Abandonment of the silver standard, that began with Britain, ultimately resulted in only a few countries left on that standard. The inherent instability of a bimetallic system is discussed well, though entirely via authority (Jevons, Royal Commission on Gold and Silver, Barbour, Rochussen). There is excellent discussion of the superseding of legal ("unrated") bimetallism with effective monometallism. The long-run decrease in the price of silver relative to gold is emphasized. Problems of the Latin Monetary Union involved paper-money issuance within the Union and overvaluation of silver that led to Holland and Germany unloading that metal onto the Union.

Third, the entry on gold points is brief but enlightening. The only limitation is the merging of coin and bullion flows, a characteristic shared by my own Gold Standard essay! Consistent with that essay, the entry recognizes that "gold movements begin before these points are reached, as some business houses with special facilities, or undertaking large transactions, find a profit in remitting gold at much closer rates" (Palgrave, 1896, p. 227).

16.2 Classical Gold Standard (Chapter 14)

The literature on the classical gold standard is voluminous and not fully encompassed by the essay (in Chapter 14) and its references. One regrettable omission is the "trilemma approach" to the gold standard or any monetary system for an open economy. Such a country has three jointly unattainable objectives: fixed (or stable) exchange rate, free international capital mobility, and autonomous (or domestically oriented) monetary policy. One objective must be sacrificed for the other two to be realized.[3] The classical gold standard is generally interpreted as dropping the last objective. Maurice Obstfeld, Jay C. Shambaugh, and Alan M. Taylor [OST] (2005) adopt the view that the U.K. ran the international orchestra, show that it did so successfully during the 1870–1913 period, and judge that "the data" (meaning their econometrics) support the conventional ("historical narratives") view: "The classical gold standard was a highly globalized period of mostly fixed rates, unfettered capital mobility, and, hence, limited monetary independence" (OST 2005, p. 424). They make the insightful statement: "Because the U.K. interest rate always resided within reasonably small bands, countries could partially adjust to the U.K. rate change and use the margin afforded by the gold points and arbitrage costs to cover the rest" (OST 2005, p. 434).

16.3 Interwar Gold Standard (Chapter 13)

The comment of insufficiency applies even more to the interwar gold standard, discussed in the gold-standard essay (Chapter 13) with only about half the space devoted to the classical gold standard and, unlike the classical standard, void of an individual essay. With *Golden Fetters*, Barry Eichengreen literally and figuratively "wrote the book" on the interwar experience. Elsewhere (Officer, 1992) I show unrestrained enthusiasm regarding Eichengreen's iconic work.[4] Appearing almost 20 years after *Golden Fetters*, Olivier Accominotti (2020) is well-worth reading as a later treatment of the interwar gold-exchange standard. And OST (2004) successfully apply the trilemma model to the interwar experience: "The trilemma was a constraint on policy for countries that fixed their exchange rate and maintained open capital markets. They lost much of their monetary autonomy compared with countries that adopted alternative regimes" (OST 2004, p. 106).

16.4 Silver Standard, Bimetallism (Chapter 13)

The Silver Standard and Bimetallism essays are only brief syntheses. In addition to the references therein (and related work published subsequently), two studies outside the mainstream warrant mention. Alejandra Irigoin (2020) provides a sympathetic account of silver-standard history as emanating from Spanish American production. Claude Diebolt and Antoine Parent (2008) argue, and present econometric evidence, that automatic "bimetallic stabilization" (described in the Bimetallism essay) was complemented with co-operative discount-rate policy of the Bank of England and Bank of France.

16.5 Bretton Woods System (Chapter 15)

It is unusual for a collection of essays to have a book review constitute an (albeit brief) chapter—but I could not resist for Chapter 15. To supplement, Michael D. Bordo (2019) and Barry Eichengreen (2013) are excellent histories.[5] OST (2005) demonstrate that the trilemma approach continues to work as an analytical device during Bretton Woods: "fixed exchange rates did not provide much of a constraint on domestic interest rates, a clear by-product of widespread capital controls…As capital controls became more porous over the 1960s, the combination of exchange rate pegs and monetary independence became untenable" (OST 2005, pp. 424, 436).

The Bretton Woods system was inextricably linked with the International Monetary Fund, and the Fund is a specialized agency of the United Nations. In Officer (2007, part II), I compare financing of the two organizations. The use of exchange rates (rather than purchasing power parity) as the ability-to-pay measure understates the GDP of less-developed countries, resulting in (desired) lower UN assessments but (undesired) lower IMF quotas. Richard Sylla (2007) relates this dichotomy to prior historical episodes in which opposite positions are taken by economic actors in light of their self-interest.

> Officer's discussion is remindful of the debates over slavery at the U.S. constitutional convention, in which the northern-state delegates argued that slaves ought to be counted for purposes of taxation but not representation, and the southern delegates argued for just the opposite, or of the debates between Britain and its colonies in the heyday of the empire, in which the British wanted the colonies to be economically independent but politically dependent, whereas the colonies wanted just the opposite.

16.6 Further Discussion

Detailed description and analysis of various monetary standards (gold, silver, bimetallic, Bretton Woods, paper) are presented for the United States and, to a lesser extent, Britain in Parts V and VI.

Notes

1. Steven N. Durlauf and Lawrence E. Blume (2008).
2. R. H. Inglis Palgrave (1894, 1896, 1899). For a fascinating history of the original *Dictionary*, see Murray Milgave (2018).
3. Therefore the trilemma is also termed "the impossible trinity" (Joshua Aizenman 2019, p. 445).
4. "Eichengreen assembles a gigantic amount of material and synthesizes a tremendous amount of literature" (Officer, 1992, p. 408).
5. My own writings on the International Monetary Fund include Officer (1978, 1990).

References

Accominotti, Olivier (2020). "International Monetary Regimes: The Interwar Gold Exchange Standard." In Stefano Battilossi, Youssef Cassis and Kazuhiko Yago, eds. *Handbook of the History of Money and Currency*, pp. 633–664. Singapore: Springer Nature.

Aizenman, Joshua (2019). "A Modern Reincarnation of Mundell-Fleming's Trilemma." *Economic Modelling* 81 (September), pp. 444–454.

Bordo, Michael D. (2019). "The Operation and Demise of the Bretton Woods System, 1958–1971." In Naomi Lamoreaux and Ian Shapiro, eds. *The Bretton Woods Agreements*, pp. 217–235. New Haven and London: Yale University Press.

Diebolt, Claude, and Antoine Parent (2008). "Bimetallism: The 'Rules of the Game.'" *Explorations in Economic History* 45 (July), pp. 288–302.

Durlauf, Steven N., and Lawrence E. Blume, eds. (2008). *The New Palgrave Dictionary of Economics*, 2nd ed. Houndmills, Basingstoke, Hampshire and New York, eight vols.

Eichengreen, Barry (2013). "The Rise and Fall of the Bretton Woods System." In Randall E. Parker and Robert M. Whaples, eds., *Routledge Handbook of Major Events in Economic History*, pp. 275–282. London and New York: Routledge.

Irigoin, Alejandra (2020). "Rise and Demise of the Global Silver Standard." In Stefano Battilossi, Youssef Cassis and Kazuhiko Yago, eds. *Handbook of the History of Money and Currency*, pp. 383–410. Singapore: Springer Nature.

Milgave, Murray (2018). "Palgrave's Dictionary of Political Economy." In *The New Palgrave Dictionary of Economics*, 2nd ed., 3rd ed. London: Palgrave Macmillan.

Obstfeld, Maurice, Jay C. Shambaugh, and Alan M. Taylor [OST]. (2004). "Monetary Sovereignty, Exchange Rates, and Capital Controls: The Trilemma in the Interwar Period." IMF *Staff Papers* 51 (Special Issue), pp. 75–108.

———. (2005). "The Trilemma in History: Tradeoffs Among Exchange Rates, Monetary Policies, and Capital Mobility." *Review of Economics and Statistics* 87 (August), pp. 423–438.

Officer, Lawrence H. (1978). "An Insider's History of the International Monetary Fund: A Review Article." *Journal of Economic Issues* 12 (September), pp. 677–686.

______ . (1990). "The International Monetary Fund." In *International Trade: The Changing Role of the United States, Proceedings of the Academy of Political Science* 37, No. 4, pp. 28–36.

———. (1992). Review of Barry Eichengreen, *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939*, in *Business History Review*, Vol. 66, No. 2 (Summer), pp. 407–409.

———. (2007). *Pricing Theory, Financing of International Organizations and Monetary History*. London and New York: Routledge.

Palgrave, R. H. Inglis, ed. (1894, 1896, 1899). *Dictionary of Political Economy*. London and New York, three vols.

Sylla, Richard (2007). Review of Lawrence H. Officer, *Pricing Theory, Financing of International Organisations and Monetary History*, in EH.net. Economic History Services.

PART V

Anglo-American Monetary Standards

CHAPTER 17

American Monetary Standard

Originally published in *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*, pp. 11–33, 286–291, copyright Cambridge University Press, 1996. Reprinted with permission.

17.1 Unit of Account

Until the mid-1790s the monetary system of colonial times continued to exist in its basic form in the United States. The most important feature of that system was the dichotomy between the medium of exchange and the unit of account. The "Spanish dollar"—a silver coin produced in Mexico and Peru as well as Spain, and therefore sometimes called the "Mexican dollar"—was the dominant coin in transactions. Known in Spanish as the peso or piastre, it was termed the "dollar" or "piece of eight" in England and the colonies. The word "dollar" is a corruption of "thaler," an abbreviation of "Joachimsthaler," a silver coin produced in 1517 in a Bohemian county of the same name.[1] The smallest subdivision of the peso was the "real," one-eighth of a dollar. The term "piece of eight" flowed naturally for the entire peso.[2]

In contrast, the unit of account was based on the English system of pounds(£), shillings (s.), and pence (d.), where £1 = 20s. and 1s. = 12d.The phrase "based on" rather than "equivalent to" is used advisedly, because "a shilling from the British mint was not a shilling in any colony" (Carothers 1930, p. 34). Instead of a national, homogeneous, standard

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_17

of value, the individual states, as the colonies before them, had their own units of account. For example, a British shilling (1 s.) was equal to 1s. 6d. in Massachusetts but 2s. in New York (Carothers 1930, pp. 34, 47; Stewart 1924, p. 19).

These properties of the colonial monetary system continued to exist after General Cornwallis surrendered at Yorktown in 1781, ending the American Revolutionary War, and even after the Articles of Confederation were superseded by the US Constitution in 1789. However, reform of the standard of value was on its way. In 1782 Robert Morris, Superintendent of Finance under the Confederation (equivalent to Secretary of the Treasury under the Constitution), submitted a report—in response to a directive from Congress but apparently already completed and by his assistant, Gouverneur Morris (no relation—that recommended a decimal system of currency, with "units" (mills), "cents" (100 units), and "marks" (1000 units). Morris understood the concept of a unit of account; for he notes that "there is no necessity that this money unit be exactly represented in coin."[3] Morris was the first author in history to suggest a monetary framework (unit of account and medium of exchange) based on a decimal system.

The monetary reform of Congressman Thomas Jefferson, probably writing in 1783, also involved a decimal system, with the dollar as the basic unit, a "tenth (of a dollar]" or "bit," forerunner of the dime, and the smallest coin a copper "hundredth [of a dollar]," what would be called the cent.[4] The Jefferson plan was embodied in a report dated May 13, 1785, of the Congressional Committee on Finance ("Report of a Grand Committee on the Money Unit"), which also allowed for a copper piece of l/200th of a dollar. The first legislation on the subject occurred on July 6, 1785, when Congress resolved that the money unit of the United States should be the dollar with a decimal system of coinage ("the several pieces shall increase in a decimal ratio") and smallest coin a copper l/200th of a dollar. It was followed by the Act of August 8, 1786, which specified "that the money of account…proceed in a decimal ratio," with mills (1000 to the money unit), dimes (ten to the unit), cents (100 to the unit), and dollars (the money unit).[5]

On April 15, 1790, the House of Representatives directed that Alexander Hamilton, Secretary of the Treasury, prepare a report on the establishment of a mint. Hamilton reported to the House on January 28, 1791, with a recommendation for a decimal system of account implicit in his coinage: the dollar, "tenth part" of the dollar, and "hundredth part" of the dollar (and also a "two-hundredth" part of the dollar).[6] Still, the standard of value remained as it had been in colonial America.

Finally, on April 2, 1792, Congress passed the Mint Act, which declared that "the money of account of the United States shall be expressed in dollars or units, dismes or tenths, cents or hundredths, and milles or thousandths." The decimal system was adopted! Further, for implementation: "all accounts in the public offices and all proceedings in the courts of the United States shall be kept and had in conformity to this regulation."[7] As commented by Stewart (1924, p. 18):

> At one fell swoop with a few chosen words the English system of accounting with pounds, shillings, and pence that had been used by the people under the Colonial government of Great Britain and continued after the Declaration of Independence, as a matter of necessity down to the passage of this momentous Act of Congress, was obliterated as far as public records were concerned.

It took until about 1800 for the private sector to follow the government and courts in moving to the uniform national decimal accounting system from their states' specific pound-shilling-pence units of account.[8]

17.2 Metallic Content of the Dollar and Coinage Denominations

It also took a long time for the United States to reform the media of exchange that it inherited from colonial times. The new nation continued to rely on foreign coin, with the Spanish dollar and its fractional parts dominant. This dollar was rated by the individual states at differing values in terms of the local unit of account (although some valuations were common to several states—Stewart 1924, p. 19). A silver standard was in effect, then, and its basis was the Spanish dollar. It is true that the dollars in circulation tended to vary greatly in weight and fineness (ratio of pure metal to total weight). This was because of the practice of sending the full-bodied coins abroad to settle balance-of-payments deficits, the lack of quality control at the Spanish mints in Mexico and Peru, and the private clipping and sweating of coins (in order to remove particles of silver prior to recirculation). Yet a dollar coin, irrespective of its condition, was acceptable "by tale" everywhere in the United States (as it had been in the colonies), that is, at its full assigned nominal value in local currency. While some gold coins did circulate, they were rated in dollars according to actual pure metal content.[9] This practice also dated from colonial times.[10]

The reform offered by Robert Morris involved a silver standard with the money unit only 0.25 grain of silver and equal to l/1440th of the Spanish dollar.[11] The largest-denomination coin, the mark, would equal 1000/1440th of the dollar. The number 1440 was selected so that the shilling of each state (excluding South Carolina, an outlier due to paper-money inflation) could be converted into the smallest whole number of units based on the state's valuation of the Spanish dollar. Thus Morris would effect a reconciliation of the Spanish dollar, the new national unit, and the individual states' existing units of account. "The Morris plan was not only ingenious, but the most cumbersome scheme for coinage ever devised by Man" (Taxay 1966, p. 16).

In contrast, Jefferson advocated as the monetary unit the dollar, equal in value to the Spanish dollar and composed of 365 grains of pure silver (purported to be the fine-metal content of the newest Spanish dollar). He would coin also a silver half-dollar and lesser denominations, along with a $10 gold piece. Thus there would be a bimetallic standard, and the gold/silver price ratio (ratio of silver to gold per dollar coinage) would be set at 15. The British fineness (ratio of pure metal to total weight) of 11/12th would be adopted. The Finance Committee report of 1785 modified Jefferson's plan, principally by having a 362 rather than 365-grain silver dollar and a $5 rather than $ 10 gold coin. It was followed by three reports on coinage and a mint, produced by the Board of Treasury, that took over the responsibilities of the Superintendent of Finance after Morris resigned. The Act of August 8,1786 was heavily influenced by these reports.[12]

On October 16, 1786, Congress passed its first bill to establish a mint, in accordance with the Act of August 8, 1786, which authorized the coinage of a silver dollar and half-dollar, and a $10 gold piece, called the eagle, and a half-eagle, plus smaller denominations. Table 17.1 assembles all legislation and practice on the gold and silver value of the dollar from the first Mint Act to 1934. The final column of Table 17.2 presents the most important element in Table 17.1, the fine-metal content of the dollar, to the present day.

The specified amount of fine (pure) metal in a coin is the product of the standard fineness and standard weight. The standard fineness is the stipulated proportion of the weight consisting of pure metal, the remainder being alloy and considered worthless. The standard weight is the specified gross weight (pure metal plus alloy). Gold or silver of standard fineness is called "standard gold" or "standard silver." It is the relative amount of fine metal in gold and silver coins for a given valuation (say, a dollar) that defines the gold/silver price ratio (called the mint ratio), which was 15.25 under the 1786 legislation. However, except for a few copper pieces produced by a private contractor, the legislation was not put into effect and the mint was not established.

Yet the impetus for change was strong. In his report of January 1791, Hamilton investigated the specie market value of the Spanish dollar in the United States. Merchants valued the dollar at 24.75 grains of fine gold, while examination of the existing dollar coins in circulation revealed a gross weight of 416 grains on average and a fine weight of 368 and 374 grains for the two most recent dollars minted. These figures yield market gold/silver price ratios of 14.87 and 15.11, leading Hamilton to suggest a mint ratio of 15, resulting in a dollar of 24.75 grains of pure gold and 371.25 grains of pure silver. With fineness of 11/12th, standard weights would be 27 grains for the gold dollar and 405 for the silver.[13] Hamilton also recommended the minting of a $10 gold piece but not a half-dollar silver coin. The reverse coinage was inconceivable. Given the high valuation of gold relative to silver, a $10 silver coin would have been much too large and a half-dollar gold coin too small.

On March 3, 1791, Congress ordered that a mint be established, and on April 2, 1792 it passed the second Mint Act in US history but the first under the Constitution. Relying heavily on Hamilton's recommendations, Congress nevertheless deviated from them in authorizing the coinage also of gold half-eagles and quarter-eagles and of silver half-dollars and quarter-dollars. Also, incredibly, it legislated a cumbersome fineness of the silver dollar, at 1485/1664th (see Table 17.1).[14]

The first coinage of the mint consisted of experimental half-dimes, probably produced in July 1792.[15] While coinage of minor copper coins began the following year, gold and silver could not be processed until the assayer and chief coiner posted a $10,000 bond, under the Mint Act, and the appointees were unable to do so. Thomas Jefferson as Secretary of State wrote to President Washington about the problem, and on March 3, 1794 Congress reduced the amount of the bond required.[16]

With the posting of bonds, the mint could effectively function. The first deposit of bullion (which was silver), however, did not take place until July 1794, and its coinage was completed on October 15.[17] This began

Table 17.1 Gold and Silver value of American Dollar, 1786–1934

Authority	*Date*	*Type of Specification*	*Gold Dollar*			*Silver Dollar*			*Gold/Silver Price Ratio*[a]
			Fineness	*Weight (grains)*		*Fineness*	*Weight (grains)*		
				Standard	*Fine*		*Standard*	*Fine*	
Congress	August 8, 1786	Legal	11/12	26.8656[b]	24.6268[b]	11/12	409.7891	375.64	15.25
Congress	April 2, 1792	Legal	11/12	27[b]	24.75[b]	1485/1664	416	371.25	15.00
US Mint	October 15, 1794[c]	Practice	11/12	27[b]	24.75[b]	9/10	416	374.4	15.13
Congress	June 28,1834[d]	Legal	116/129	25.8[b]	23.2[b]	1485/1664	416	371.25	16.00
Congress	January 18, 1837	Legal	9/10	25.8[b]	23.22[b]	9/10	412.5	371.25	15.99
Congress	February 12, 1873	Legal	9/10	25.8	23.22	–	–	–	–
Congress	February 28, 1878	Legal				9/10	412.5	371.25	15.99
Congress	March 14,1900	Legal	9/10	25.8	23.22				
President	January 31, 1934	Legal	9/10	15 5/21	13 5/7				

[a]Ratio of fine-silver to fine-gold content of dollar
[b]Weight of gold dollar is 1/10 of eagle ($10 gold piece)
[c]Terminated October 28, 1795
[d]Effective July 31,1834

Sources Huntington and Mawhinney (1910, pp. 475, 496–7, 502, 534–5, 579, 610), *International Monetary Conference* (1879, p. 450), Krooss (1969, p. 2805), Select Committee on Coins (1832, p. 17)
Congress = Act of Congress
President = Presidential Proclamation

an amazing episode in American monetary history. For over a year—overlapping the tenure of two directors of the mint—silver was coined at a convenient fineness of 9/10th rather than the slightly lower fineness of 1485/1664th specified by law (see Table 17.1). The practice was quite deliberate, though with the expectation that the law would be changed to correspond to mint practice. Instead, in the short run, the mint practice changed to correspond with the existing law; for, when the third mint director took office on October 28, 1795, he ordered that the legislated fineness be followed.[18] However, 42 years later, the law would be amended to conform to the 1794–1795 mint practice.

On June 28, 1834 Congress passed legislation (effective July 31) that drastically changed the mint ratio, from 15 to slightly over 16. The silver dollar was left unchanged, but the standard and fine weights of the gold dollar were reduced non-proportionately, resulting in an unwieldy fineness of the gold dollar (I16/129th) to accompany that of the silver dollar (1485/1664th). These deficiencies were removed with the legislation of January 18, 1837, when the fineness of both gold and silver coins was changed to 9/10th, the former by increasing the fine metal in a dollar's worth of coin, the latter by reducing the standard weight of a silver dollar. The mint ratio now moved slightly below 16.[19]

The *mint price* of gold or silver is the value of domestic money that the mint will coin per physical unit of bullion deposited with it. The standard weight of the gold dollar established in the legislation of 1837 was to remain until 1934 (see Table 17.1). From the standard weight, one can obtain the mint price of standard gold: the price per "standard ounce" (meaning ounce of standard fineness) at which bullion is converted into coin at the mint. From 1837 to 1934, the mint price of standard gold was \$18.604651+ per ounce, obtained as the ratio 480/25.8, where the numerator is the number of grains per ounce, the denominator the standard weight of the dollar in grains, and " + " represents additional decimal places. Taking 10/9th of that price, where 10/9th is the reciprocal of the standard fineness, the mint-price equivalent per ounce of fine gold was \$20.671834+.

The Act of March 3, 1849 authorized coinage of gold dollars and double-eagles (\$20 coins). The former was not popular, in part because of its small size, in part because of the introduction of subsidiary silver coinage in 1853 (see Sect. 17.3), with less than \$20 million produced in total until the coin was discontinued by the Act of September 26, 1890. The Act of February 21, 1853 allowed for coinage of a \$3 gold piece, an unusual denomination that also was discontinued by the 1890 legislation, with less than $2 million having been struck. In contrast, the double-eagle was enormously successful, and the preferred coin of those engaged in international gold operations (see Officer 1996, Sect. 2.2.1(2) of chapter 9); almost $3.5 billion of this coin was produced in total. The Act of February 12, 1873 summarized the coinage authorizations and standards for gold. Both this act and that of March 14, 1900 are distinguished by defining the dollar directly in terms of gold, whereas all earlier legislation had established the weight of the gold dollar implicitly via the eagle.[20].

Even World War I did not disturb the established weight of the gold dollar. However, the Act of May 12, 1933 authorized the President to reduce the gold content of the dollar to a minimum of 50% of its existing weight. Subsequently, the Gold Reserve Act of January 30, 1934 amended that provision by authorizing the President by proclamation to fix the weight of the dollar at any level between 50 and 60% of its current weight. On January 31, 1934 President Roosevelt reduced the weight of the dollar to 59.06% of its existing level. With the 9/10th fineness retained, the fine weight of the dollar became 13.7142+ (13 5/7) grains of gold and the "mint-price equivalent" 480/(13.7142+) = $35 per fine ounce.[21]

Turning to Table 17.2, the Smithsonian Agreement of December 18, 1971 increased the price of pure gold to $38 an ounce, implying a fine gold weight of the dollar equal to 480/38 = 12.6315+ (12 12/19) grains. The final official action on the dollar/gold price occurred on February 13, 1973, when the dollar was devalued to $42.22… per ounce, equivalent to 480/42.22 … = 11.3684+ grains of pure gold in a dollar (where "…" indicates an infinitely recurring pair of numbers, in this case "22").

17.3 Legal-Tender Status of Coin

The Mint Act of 1792 established full legal tender for all gold and silver coins issued by the mint, those of less than full weight at values proportional to their weight. With the other prerequisites of a specie standard satisfied (see Officer 1996, Sect. 2 of chapter 2), bimetallism was legally installed. The Act of 1837 declared legal tender at full nominal value, with no reduced value for lightweight coin—a status received also by the new denominations of gold coin in the Acts of 1849 and 1853.

Table 17.2 Effective American monetary standards, 1791–present

Time Period	*National Standard*	*Exceptional Standard*	*Fine Metal in Dollar*[a] *(grains)*
January 1, 1791–March 2, 1794	Silver	–	371[b]
March 3, 1794–October 14, 1794	Silver	–	371.25
October 15, 1794–October 28, 1795	Silver	–	374.4
October 29,1795–August 29, 1814	Silver	–	371.25
August 30, 1814–February 19, 1817	Paper	silver (New England)[c]	371.25
February 20, 1817–July 30, 1834	Silver	–	371.25
July 31, 1834–January 17, 1837	Gold	–	23.2
January 18, 1837–May 9, 1837	Gold	–	23.22
May 10, 1837–May 9, 1838	Paper	–	23.22
May 10, 1838–October 9, 1839	Gold	paper (Philadelphia)[d]	23.22
October 10, 1839–March 17, 1842	Paper	gold (New England and New York)[e]	23.22
March 18, 1842–October 13, 1857	Gold	–	23.22
October 14, 1857[f]–December 13, 1857	Paper	gold (Alabama, Kentucky, and New Orleans)	23.22
December 18, 1857 [g]–December 29, 1861	Gold	paper (Philadelphia, Baltimore, and South)[h]	23.22
December 30, 1861–September 24, 1873	Paper	gold (West Coast)[i]	23.22
September 25, 1873–October 22, 1873	Currency[j]	gold (West Coast)	23.22
October 23, 1873–December 31, 1878	Paper	gold (West Coast)	23.22
January 1, 1879–August 2, 1893	Gold	–	23.22
August 3, 1893–September 2, 1893	Gold[k]	–	23.22
September 3, 1893–October 30, 1907	Gold	–	23.22
October 31. 1907–December 30, 1907	Gold[k]	–	23.22
December 31, 1907–April 5, 1917	Gold	paper (New York)[l]	23.22
April 6, 1917–March 17, 1922	Paper	–	23.22

(continued)

Table 17.2 (continued)

Time Period	*National Standard*	*Exceptional Standard*	*Fine Metal in Dollar*[a] *(grains)*
March 18, 1922–March 5, 1933	Gold	–	23.22
March 6, 1933–January 30, 1934	Paper	–	23.22
January 31,1934–August 14,1971	Paper	gold (official foreigners)[m]	13 5/7
August 15,1971–December 17, 1971	Paper	–	13 5/7
December 18, 1971–February 12, 1973	Paper	gold (official foreigners)[m]	12 12/19
February 13, 1973–present	Paper	–	11.3684

[a]Metallic dollar, even during periods of paper standard
[b]Average weight of two Spanish-dollar issues in circulation, as computed by Alexander Hamilton
[c]Also, silver standard continued in Ohio and Kentucky until January 1, 1815, and in Nashville, Tennessee, until August 1815
[d]Until August 13,1838. Also, banks in the South and West
[e]Also, in Pennsylvania January 15 - February 3, 1841, and elsewhere around that time
[f]Date pertains to New York. September 24–26 in Baltimore and Philadelphia. A large part of the banks in Pennsylvania, Virginia, Maryland, and Rhode Island also suspended before New York
[g]Date pertains to New York. Baltimore and Philadelphia returned to gold February 3–5, Virginia May 1, 1858
[h]Virginia, Baltimore, and Philadelphia banks suspended November 20–22, 1860, followed by St. Louis, South Carolina, and Georgia banks November 28–30. Philadelphia and Baltimore banks resumed February–March 1861
[i]Also, Ohio, Indiana, and Kentucky banks did not suspend until January 1861, March 1861, and March 1862, respectively
[j]Notes only, excluding certified checks
[k]Obtainable at par from currency only, not certified checks
[l]August-October 1914
[m]US Treasury bought gold from anyone but sold it only to foreign monetary authorities and licensed industrial users

Sources Bolles (1894, vol. II, pp. 264–5, 322, 328), Brown (1940, p. 37), Calmoris and Schweikart (1991, pp. 822–3), Chandler (1958, pp. 103–4), de Vries (1985, vol. I, pp. 66, 69), Dunbar (1904, pp. 280–4, 288, 309–10), Friedman and Schwartz (1963, pp. 7, 27, 471), Hammond (1957, pp. 227, 246–8, 481), Knox (1900, pp. 76, 504–5, 512, 514), McGrane (1965, p. 201), Mitchell (1903, pp. 40–1,141, 144–7), Nussbaum (1957, pp. 174–80), Secretary of the Treasury (1830, p. 47; 1838, p. 5; 1920, pp. 181–2; 1922, p. 72), Smith and Cole (1935, pp. 26–9), Sprague (1910, pp. 57, 67,187, 280–2; 1915, p. 528), Sumner (1874, pp. 68, 74, 140, 144–6, 151–2), Temin (1969, p. 113), Yeager (1976, pp. 65, 580–1). Also, Table 17.1 and see text

In the meantime the legal-tender status of foreign coin, particularly the Spanish dollar, required attention. After all, foreign coin had been the only coin of consequence in the colonies and the United States until the US mint began full functioning. Also, Hamilton had recommended gold and silver values of the US dollar derived from market values of the Spanish dollar, and the 1792 Mint Act had declared the value of the US silver dollar to be equal to that of the then current Spanish dollar.

Hamilton suggested a maximum three-year period for circulation of foreign coin. One month after the Mint Act of 1792 was passed, a committee was appointed to consider the role of foreign coin, leading to the Act of February 9,1793, which provided for the full legal tender of (1) gold coins of certain countries at stipulated valuations in proportion to their weights; (2) the Spanish silver dollar, if minimum weight of 415 grains, at 100 cents (that is, equal to the US dollar), and in proportion for parts of the dollar; and (3) the French silver crown, similarly treated. Three years after the mint began coinage, the date to be proclaimed by the President, all foreign coins except the Spanish dollar and its subdivisions would cease to be legal tender. The proclamation was made by President John Adams on July 22, 1797; but the termination date of October 15, 1797 did not stand, and various Acts between 1798 and 1834 continued the legal-tender status of foreign coin.[22]

The legislation addressing the legal-tender position of foreign coin was confusing and unhelpful to US monetary development. Foreign gold coins, no matter how worn, possessed full legal tender; but the dominant foreign coin, the Spanish silver dollar, was required to be at least 415 grains (with parts in proportion) to have this power. "The impossibility of weighing coins in retail trade meant that the entire mass of Spanish coins would be accepted as legal coins" (Carothers 1930, p. 67). The provisions for legal tender of parts of the Spanish, later Mexican, silver dollar were inconsistent and unclear. Irrespective of legislation, the public apparently considered all foreign coin to be full legal tender. In fact, until July 31, 1834, with the coming into effect of the 1834 mint legislation, the Spanish dollar was the dominant coin in the United States (see Sect. 17.8 below and Officer 1996, Sect. 3 of chapter 5 and Sect. 2.2.1 of chapter 9), and only in 1853, when US silver coins below a dollar became subsidiary in nature (see below), did the parts of the Spanish dollar lose substantial circulatory power. Finally, the Act of February 21, 1857 terminated the legal-tender status of all foreign coin.[23]

As mentioned, the legal position of silver coins changed in 1853, when an Act of February 21 provided for fiduciary coinage of silver pieces below a dollar, reducing their weight (but not fineness) by 6.91% and limiting their legal-tender power to $5.[24] It appears obvious that the purpose of the Act was to render a proper subsidiary coinage, and indeed that was its effect. Nevertheless, some observers see the legislation as paving the way to formal gold monometallism, even though the silver dollar coin is not even mentioned in the Act.

Watson (1899, p. 107) writes: "When this bill was being discussed in Congress, it was claimed that it would result in establishing the gold standard, and it clearly appears from the debates as published in the *Congressional Globe* that such was the evident purpose of the bill, and that there was no desire to conceal it." Laughlin (1900, p. 92) declares: "This, then, was the act which really excluded silver dollars from our currency." Myers (1970, p. 34) comments: "The debates in Congress show quite clearly that the intention of the Act of 1853 was abandonment of the bimetallic standard."

Rather, Carothers (1930, pp. 120, 136) is on the mark in stating: "it was not the intention of Congress to demonetize silver, and the law did not effect this result…the subsidiary coinage law of 1853 did not establish the gold standard and was not intended to establish it." But he does concede that "it paved the way to the gold standard." Martin (1973, p. 841) notes: "Although the Act of February 21, 1853 did not repeal completely the de jure bases of bimetallism, it did terminate de facto bimetallism in the United States."

Twenty years later, the Act of February 12, 1873 ended coinage of the silver dollar. A "trade dollar," of 420 grains, was to be coined freely (that is, from any depositor of silver bullion), but it was included with subsidiary coins and given the $5 legal-tender restriction. While the standard silver dollar was no longer to be issued, the full legal-tender power of existing silver dollars remained undisturbed. Legislation of June 22, 1874 revised the statutes so that [all] silver coins of the United States should have limited tender of $5 in any payment.[25] It was this 1874 action that accomplished the true demonetization of silver.

Nevertheless, it was the 1873 exclusion of silver dollars from coinage that is viewed as the formal end of bimetallism in the United States.[26] The Act of February 12, 1873 became known as the "Crime of 1873," and silver coinage became a domestic political issue.[27] Friedman (1990b, pp. 1165–6) shows that there was no crime in the legal sense, but the standard silver dollar was omitted from the coinage list intentionally, deliberately to demonetize silver. The most incisive comment on the matter is made by Watson (1899, p. 119): "no one seemed to have discovered that the Act of 1873 omitted the silver dollar until some years later when the price of silver bullion began to fall. Then the agitation began in the silver States about the omission of the dollar from the act, and it was charged that a crime had been committed."

It is ironic that the move from formal bimetallism to a legal gold standard, whether in 1873 or 1874, occurred during the greenback, *paper* standard, with a dual legal and effective gold standard not reached until 1879 (see Sects. 17.4 and 17.9). Meanwhile, silver legislation proceeded. The Act of July 22, 1876 ended both the legal-tender power and the free coinage of the trade dollar. In fact, the Secretary of the Treasury terminated coinage of the trade dollar in 1878, except for trivial amounts over the next five years. Less than $36 million had been coined in total.

Free coinage of silver was never restored. The Bland-Allison Act of February 28, 1878 directed the Secretary of the Treasury to purchase $2–4 million of silver bullion monthly, to be coined into standard silver dollars of unlimited legal tender. In fact, the Treasury purchased silver in minimum amounts, and "the coins were very badly received" (Carothers 1930, pp. 282–3). The Sherman Purchase Act of July 14, 1890 changed the monthly amount of silver bought to 4.5 million ounces, with payment to be made by the issuance of Treasury notes (see Sect. 17.4). The Act of November 1, 1893 repealed these provisions, but coinage of silver dollars continued to 1904.[28]

The legal gold standard was established in the Act of February 12, 1873, which stipulated that the gold dollar of standard weight "shall be the unit of value." Consistent with the 1792 Act, US gold coins were to be legal tender in all payments at nominal value when not below standard weight and limit of tolerance, otherwise at valuation in proportion to their actual weight. The Gold Standard Act of March 14, 1900 declared the gold dollar of 25.8 grains 9/10ths fine to be "the standard unit of value." It was not until January 30, 1934, with the passage of the Gold Reserve Act, that the gold standard was legally terminated. All gold coin was to be withdrawn from circulation and no further gold coining was to occur.

17.4 Convertibility of Government/Central-Bank Paper into Coin

The first US-government paper currency convertible into coin on demand was the Treasury demand notes authorized by the Acts of July 17,1861 and February 12, 1862. By the Act of March 17, 1862 these notes were made receivable for all payments due to the United States and for all claims and demand against the United States except for interest on bonds and notes, which were to be paid in coin. Also, they were made legal tender for all payments (except for interest payable by the United States, as aforementioned).

Four Acts (February 25 and July 11,1862; January 17 and March 3, 1863) authorized the Treasury to issue United States notes, the famous "greenbacks". These legislations are called the "legal-tender acts," because, for the first time, they gave legal-tender power to paper currency (US notes and Treasury demand notes). The legal tender of US notes excluded both payment of import tariffs and interest on the public debt. The notes were not redeemable in coin until January 1, 1879, as provided by the Resumption Act of January 14, 1875.[29] The Gold Standard Act of March 14, 1900 repeated this obligation of the Treasury, and stipulated gold coin of the weight and fineness fixed in the Act, which was the standard weight and fineness.

The Act of March 3, 1863 provided a third government paper currency: gold certificates, issued in response to gold coin and bullion deposited with the Treasury and payable in such gold on demand. The amount issued could be up to 20% higher than the value of gold deposited. By the Act of July 12, 1882, (1) only gold coin was so depositable and payable, (2) the certificates were made receivable for customs, taxes, and all public dues, and (3) the certificates were to be pure warehouse receipts, with no excess of issuance over deposits.[30] The Act of March 2, 1911 extended the allowable deposits to foreign gold coin and gold bars produced by US mints or the New York assay office. The certificates were a convenience to those dealing in gold. It was not until the Act of December 24, 1919 that they were made legal tender.[31]

The fourth government paper currency was the Treasury notes of 1890. The Act of July 14, 1890 established that these notes were to be a full legal tender (except where otherwise expressly stipulated by contract) and receivable for customs, taxes, and all public dues. The notes were made redeemable on demand, in gold or silver coin at the discretion of the Treasury. The Act of March 14, 1900 specified gold coin, of standard weight and fineness (as for US notes).

Federal Reserve notes, authorized by the Federal Reserve Act of December 23, 1913, were a central-bank currency to be redeemed in gold at the Treasury on demand. The notes were made receivable by all Federal Reserve banks, member banks of the Federal Reserve System, and national banks, and also for all taxes, customs, and other public dues. However, they were not given legal-tender status until June 5, 1933.[32]

17.5 Nature of Coinage

Coinage of private bullion has three characteristics: openness (delineation of the depositors for whom coinage would be provided), cost (with charges possible for mint expenses and for seigniorage, the monopoly profit of the mint), and speed (the duration between receipt of bullion and delivery of coin to the depositor).

Regarding openness, the Mint Act of 1792 specified free coinage (open to bullion from any person or persons) and for both gold and silver, while that of 1873 declared free coinage only for gold (with a $100 minimum) and silver trade dollars. The Federal Reserve Act of 1913 permitted Federal Reserve banks to deal in gold coin or bullion and to exchange Federal Reserve notes for gold, gold coin, or gold certificates. This provision meant, in particular, that the owner of gold bars or foreign gold coin could always convert it into American gold coin at the mint or, if a Federal Reserve bank was prepared to transact, could exchange the bullion for gold-convertible Federal Reserve notes. The Federal Reserve Bank of New York did exercise its right to purchase gold bars, and from all comers, at least from 1925 onward. Even throughout World War I, the Treasury continued its open policy of buying gold according to statute.

Turning to the mint charge, Robert Morris suggested that it be almost 3.5%, while the Acts of August 8 and October 16, 1786 set it at almost 2% for silver and slightly above that for gold.[33] The Mint Act of 1792 involved no mint charges except if the depositor and mint were to agree to an immediate exchange of coins for standard bullion, in which case there would be a charge of 0.5%. This provision was pursuant to a recommendation in the Hamilton report. The Act of June 28, 1834 repeated the 0.5% charge, but changed "immediate" exchange to payment within five or forty days, in contradictory sentences within the same section of the Act.[34]

Neither the 1792 nor 1834 Acts assessed the depositor for the cost of mint procedures to bring deposited bullion to standard: melting (required for various other procedures), assaying (determination of the fine-gold and/or fine-silver content of bullion), alloying (addition of alloy to reduce the fineness of overly fine bullion to standard), parting or separating (separation of gold from silver in bullion containing both in significant quantity), toughening (removal of metals unacceptable for coining that are intermixed with bullion), and refining (the specific processes by which parting and/or toughening take place).[35]

The Mint Act of 1837 marked an abrupt shift in the nature of mint charges. The 0.5% charge was dropped, and the only charges permitted were for separating, toughening, refining, and for the metal used for alloy, the rates to be fixed from time to time so as not to exceed the actual expense of the mint. However, the Act of February 21, 1853 added seigniorage of 0.5%, with no reciprocation of quick coinage. The Mint Act of 1873 reduced seigniorage to 0.2% (for converting standard gold bullion into coin), and the Resumption Act of 1875 eliminated seigniorage. From that time on, coinage was gratuitous (meaning no seigniorage charge for standard bullion), but the Act of 1873 specified charges for melting, refining, toughening, and copper alloy, to be fixed from time to time, at actual average cost.

The speed of coinage was dismal until the second half of the nineteenth century. In 1803 the annual report of the mint stated that the certificates for deposits of bullion were sold to the banks at 0.25–0.5% discount for delay of coinage (Bolles 1894, vol. II, p. 165, n. 4), and a 0.5% discount for about this time is noted by Stewart (1924, p. 50). In 1831 the delay in coinage was said to be two months, equivalent to an interest loss of 1% (Sumner 1874, pp. 104–5). As late as 1850, a lag of 52 days between deposit and coinage was experienced, equal to nearly a 1% loss of interest (Committee on Commerce 1850, p. 4). An additional delay and cost emanated from the location of the mint in Philadelphia, whereas the international commercial center of the country had become New York City and there was no branch mint there. In 1850 the transport cost for shipment of bullion from New York to Philadelphia was reported at 0.25%.

It is not surprising, then, that the mint was little used by private parties. For enhanced business, two reforms were needed. First, funds had to be appropriated by Congress for the mint for the purchase of bullion in advance of deposits. Without such a “bullion fund,” the provisions of the 1792 and 1834 Acts for speedy exchange of coin for bullion were inoperative: "the depositor of bullion had to wait weeks and even months for his coins" (Carothers 1930, p. 73). As early as 1797 the Committee of Congress on the mint recommended a bullion fund (Stewart 1924, p. 50), and in 1836 the Secretary of the Treasury requested that Congress authorize him to establish such a fund, even temporarily, in the amount of $100,000 (Secretary of the Treasury, 1836, p. 1). The Mint Act of 1837 directed the Secretary to keep in the mint a bullion fund of up to $1 million (when the state of the Treasury so permitted) for the purpose of paying depositors as soon as practicable after the value of bullion had been ascertained. The Act of May 23, 1850 permitted the President to transfer funds to the mint for the same purpose (Huntington and Mawhinney 1910, p. 509). A bullion fund was in fact established (Bolles 1894, vol. II, p. 514).

The second required reform was the institution of a mint branch or equivalent office in New York. The Act of March 3, 1853 provided for the establishment of an "assay office" in that city (Huntington and Mawhinney 1910, pp. 514–16). To private parties in New York, dealing with the assay office was the same as dealing with the mint, except that the element of distance and associated expense was eliminated.[36] The New York Assay Office opened for business in October 1854, and from that date dealing with the mint (via the Assay Office) was a practicable opportunity for private parties centered in New York.

The Federal Reserve Act gave the Federal Reserve banks the right to deal in gold coin and bullion and, in particular, to exchange Federal Reserve notes for gold bars. At least in the 1925–1931 period, the Federal Reserve Bank of New York purchased bars from private parties at the mint price (see Officer 1996, Sect. 2.2.2 of chapter 9). So owners of gold bars could receive for them either coin at the New York Assay Office or Federal Reserve notes (exchangeable into coin) at the Federal Reserve Bank of New York.

17.6 Provision of Bars

The Act of May 26, 1882 authorized the mints and the Assay Office in New York to provide gold bars in exchange for US gold coin, with a $5000 minimum. The Act of March 3, 1891 specified that this exchange could occur only with the approval of the Secretary of the Treasury, and allowed the Secretary to impose a charge equal to the cost of manufacturing the bars. Finally, the Act of March 3, 1901 allowed a charge of any amount.[37]

On June 1, 1882, the 1882 Act went into effect (*New York Times*, July 2, 1882, p. 9). The bars provided by the Treasury were much prized, because they were "Assay bars," that is, bars with the fineness (also weight and value) stamp of the New York Assay Office. From March 1928, if not earlier, the bars were obtainable also from the Federal Reserve Bank of New York (see Officer 1996, Sect. 2.2.1 of chapter 9).

17.7 Convertibility of Banknotes and Bank Deposits

Though the monetary system of the American colonies extended into the United States, indeed as far forward as the 1790s (see Sects. 17.1 and 17.2), there was one respect in which the US system differed. Commercial banks did not exist in the colonies. During the Confederation period (1781–1788), three banks were established in the United States. By the end of 1790 four banks were open for business, and by the end of 1791, six. Then the growth of banking accelerated. In 1800 29 banks were operating; in 1816, 246.[38]

Historically, banks were of four types: federal, private, state, and national. Some early banks (Bank of North America, First and Second Bank of the United States) were federal in the senses that enabling legislation was by Act of Congress and the banks performed some functions of a central bank. Private banks were unchartered and unincorporated. State banks were chartered by individual states, and national banks by Congressional Acts of February 25, 1863 and June 3, 1864.[39] In 1913 the Federal Reserve Act created a true central-banking system, the Federal Reserve banks.

Because the notes issued and deposits created by banks were debts of the institution, the banks had the legal obligation of extinguishing these debts in legal tender, at par (meaning at face value, without discount) and, by contractual obligation (for notes and demand deposits), on demand. Until 1862, the only legal tender was gold and silver coin; so banks had to redeem their notes and deposits in that medium. Then there arose the concept of "lawful money," meaning money usable as legal bank reserves against note and deposit liabilities. In addition to legal-tender coin, lawful-money status was extended to US notes (Act of February 25, 1862), Treasury demand notes (Act of March 17, 1862), gold and silver certificates (Act of July 12, 1882), and Treasury notes of 1890 (Act of July 14, 1890).[40].

By the enabling legislations, the legal reserves of national banks could consist only of lawful money, and Federal Reserve notes were redeemable at the Federal Reserve banks in lawful money. While national banknotes were redeemable in lawful money at the issuing bank, the Act of June 20, 1874 provided also for their redemption, in US notes, at the Treasury, based on a 5% redemption fund (in lawful money) maintained there by the issuing banks (Huntington and Mawhinney, 1910, pp. 418–21). For state banks, acceptable legal-reserves media were determined by state legislation. In most states, not only lawful money but also national banknotes were permissible reserves (Friedman and Schwartz, 1963, pp. 21, 781–2). In sum, all banks stood ready to cash their deposits and banknotes at par in coin or in paper currency redeemable in coin.

In addition to Federal Reserve notes, the Federal Reserve Act authorized circulating notes of individual Federal Reserve banks (known as Federal Reserve banknotes). They were similar to national banknotes, but never replaced them and were not a popular currency. Although made legal tender (with Federal Reserve notes) on June 5, 1933, their issuance was terminated by the Act of June 12, 1945.[41]

17.8 Legal Versus Effective Monetary Standards

With no restrictions on the melting or use of gold and silver, all the domestic conditions for a metallic standard in general and for a specie standard in particular were legally fulfilled in the United States from 1786 onward. The one exception was a functioning mint to convert bullion into coin, which did not begin until March 1794. The lack of a mint also meant that the monetary legislation of 1786 was inoperative.

Formally, the United States was on a bimetallic standard from 1786 to 1873. However, although the Congressional legislations of 1786, 1792, 1834, and 1837 all involved a bimetallic standard, it happened that even under these laws the United States was in fact either on a gold or a silver standard but never both. The reason is that a mint gold/silver price ratio different from the market (world) ratio provides incentive for (1) the undervalued metal, whether bullion or coin, to be sold on the world market, and (2) the overvalued metal to be coined and utilized as domestic money. Indeed, if the undervalued metal is exchanged for the overvalued one, a process of arbitrage occurs that in principle can change the market ratio so that it comes sufficiently close to the mint ratio to eliminate the incentive for the transactions. In practice, there is generally insufficient supply of the undervalued metal available domestically to alter the world ratio significantly.[42]

The "world" market gold/silver price ratio for 1791–1834 is shown in Officer 1996, column 2 of table 5.2, which may be compared to the American ratio listed in the last column of Table 17.1.[43] It was only from March 1794 that the United States possessed a functioning mint to support its legislation. So the Mint Act of 1786 was irrelevant and that of 1792 applicable only from March 1794. Until that last date, a silver standard based on the Spanish dollar reigned in the United States by default (see Sect. 17.2). From March 1794 until July 30, 1834, the American legal mint ratio, at 15—and even the slightly higher unauthorized ratio in 1794–1795—was continuously below the world ratio. This meant that gold was undervalued and silver overvalued in the United States relative to world markets, and an effective silver standard resulted, notwithstanding legal bimetallism supported by a functioning mint. Interestingly, the result was unintentional. Hamilton had recommended a mint ratio that he thought was close to market rates both in the United States and abroad, and in this he was at least temporarily successful.[44]

The situation was reversed in 1834, when Congress deliberately established a mint ratio (at 16.00) above the world ratio (15.73), so that gold would be overvalued and silver undervalued. From 1834 to 1873 the world gold/silver price ratio was consistently below 16. Beginning July 31, 1834, then, it was economically unsound for a private party to provide silver for coinage and economically sound for the party to withdraw from circulation any silver previously coined, melt it down for its metal content, sell it for gold on the open market, and present the gold to the US Mint for coinage. The only metallic standard that could effectively exist and persist in the United States was the gold standard.[45] The fact that the world gold/silver price ratio rose above 16 in 1874 is irrelevant, because in 1873–1874 silver was legally reduced to a subsidiary coinage (see Sect. 17.3) and in any event the United States was then on the paper greenback standard.[46] Five years later the United States was on not only a legal but also an effective monometallic gold standard. In summary, disregarding episodic paper standards (and eventually a lasting one), the United States was on a silver standard until July 30, 1834 and on a gold standard thereafter.

17.9 Paper Standards

Table 17.2 shows the time periods of effective monetary standards of the United States from 1791 to the present. A metallic standard, gold or silver, was subject to interruption by a paper standard. In each case the movement off specie was initiated by commercial banks "suspending specie payments" in the face of an experienced or feared increased demand for specie on the part of the public, a demand that could not be satisfied given the fractional-reserve system under which banks operated. One scenario was for the banks to run out of reserves or virtually so, to close their doors, and to declare bankruptcy. The alternative, and preferred, scenario was for the banks to refuse to pay specie for their outstanding notes and deposits at the par (dollar-for-dollar) value but nevertheless to remain open and sometimes even to expand their note and deposit liabilities. As Temin (1969, p. 115) observes: "Suspension as practiced in the nineteenth century was not bankruptcy; one might say it was an alternative to bankruptcy." For a bank to refuse to convert its note and deposit liabilities into legal tender (or, later, lawful-reserve money for national banks and legislated acceptable money for state banks) was not only in general violation of a contract but also specifically illegal in many states. However, "the laws on this matter were seldom enforced" (Temin 1969, p. 114).

With "suspension of specie payments," markets developed in which the notes of suspending banks traded at a discount in terms of specie. The existence of such markets meant that Gresham's Law was inoperative. Specie and its equivalent (notes and deposits of non-suspending institutions) circulated together with notes and deposits of the suspending institutions, but a fixed exchange rate ("parity") between the two types of money was not imposed.[47]

From 1861, with the issuance and circulation of government currency, followed by national banknotes and, after the Federal Reserve Act, central-bank currency, inconvertibility involved not only commercial-bank but also government behavior. For the government to cease to honor its redemption commitments could make the banks declare suspension. It was also true that the government standing fast could induce the commercial banks to avoid suspension. However, the existence of currency issues directly or indirectly government-guaranteed could foster banking panics by providing alternatives beyond specie for local banknotes and deposits.[48]

Suspension of specie payments did not generally occur in all regions of the country: it was not a national phenomenon. In the affected states, a floating exchange rate between (1) the paper dollar (note or deposit) and (2) specie and foreign exchange resulted. In the states where banks continued to honor the specie-convertibility commitment, the specie standard was disturbed neither within regions nor internationally but only against regions where suspension occurred. The floating exchange rate led to a currency depreciation only in those areas where banks suspended payments.

As Table 17.2 shows, widespread bank suspensions occurred in 1814–1817, 1837–1842 (accurately characterized as a series of suspensions), 1857, and 1861–1878, with localized suspensions in 1860–1861.[49] The experience of December 30, 1861–December 31, 1878 is noteworthy for several reasons. First, it was the first occasion in which not only the banks but also the government suspended; for on December 30, 1861 the Treasury refused to honor the right of holders of its demand notes to redeem the notes in gold. Second, this episode involved government issuance of legal-tender paper currency, US notes or "greenbacks," whence the name "greenback period." As the Treasury demand notes were received by the Treasury for customs duties (a property that US notes lacked), they were replaced by greenbacks in accordance with the greenback legislation. Third, a free market for gold not only developed in terms of the irredeemable dollar (Treasury demand notes and banknotes, and later greenbacks) but also was institutionalized, with a formal gold market in New York City and banks offering both gold deposits (that is, deposits payable in gold) and ordinary (greenback) deposits.[50]

In the postbellum period to World War I, the experienced bank suspensions beyond the greenback period involved the withholding of the obligation to convert deposits into both currency and specie rather than specie alone. In this respect the 1873 experience has been called a "suspension within a suspension" (Martin 1898, p. 40), as specie payments had been suspended since the end of 1861 and now banks refused to cash their deposits into currency as well. Suspension of currency (and specie) payments also occurred in 1893 and 1907. In these three episodes a premium on currency in terms of certified checks developed. The currency-premium experiences involved an appreciation of currency against deposits (certified checks), but not a depreciation of currency against gold and foreign exchange (except as already was occurring during the greenback period).[51]

In April 1917, when the United States entered as a belligerent in World War I, another paper-standard period ensued. The Treasury and the Federal Reserve banks, though always claiming that they obeyed the statutes guaranteeing redemption in gold coin of all US paper money and Federal Reserve notes on demand, in fact imposed effective, though informal, restrictions on redemption from April 6, 1917 to March 17, 1922.[52] Commercial banks cooperated with the Federal Reserve by converting their notes and deposits only into currency and not gold, a perfectly legal restriction (Cross 1923, p. 377).

The final, and still current, US paper standard began on March 6, 1933, when President Roosevelt suspended gold redemption and prohibited banks from paying out gold. Subsequently, an Executive Order of March 10 prohibited gold payments by banks and non-banks unless licensed by the Secretary of the Treasury. The Congressional Joint Resolution of June 5, 1933 declared "gold clauses"—provisions for payment in gold or in US money measured in gold—to be invalid in the sense that such obligations were dischargeable in any legal tender, and it conveyed legal-tender status on all coins and currencies of the United States (including Federal Reserve notes).

From September 8, 1933, the official gold price was fixed daily at the world market price less shipping and insurance cost, but only for the purpose of the Treasury purchasing gold from domestic mining companies. This set the stage for the Gold Reserve Act of January 30, 1934 and the Presidential Proclamation the next day that established a fixed dollar price for gold at $35 per fine ounce. However, the existing paper standard was not disturbed in any real sense. Indeed, the Gold Reserve Act forbade redemption of any US currency in gold.

Rather, from January 31, 1934 the United States was only on a "limited gold-bullion standard" (Yeager 1976, p. 65, n. 12). The Treasury purchased gold bars from all comers at $35 minus 0.25% ($34.9125) and sold them to foreign monetary authorities and licensed industrial users, *but to no one else*, at $35 plus 0.25% ($35.0875) per fine ounce. "Rather than being the basis of the monetary system …[gold became] a commodity whose price is officially supported" (Friedman and Schwartz 1963, p. 472).[53] Treasury sales of gold were suspended by President Nixon from August 15 to December 17, 1971, and were terminated on February 13, 1973. The official gold price became irrelevant.

All paper standards except the last, that beginning in 1933, were generally considered by contemporaries as temporary aberrations from the previously applicable metallic standard. In fact, in every case (except 1933) the paper standard eventually came to an end, whereupon the existing metallic value of the dollar regained its effectiveness. The paper standards differed in their effect on the foreign-exchange market. Some periods involved noticeable disturbances to the market: 1814–1817, 1837–1842, 1861–1878. Bank suspensions in other periods, especially later in the century, did not significantly affect the local foreign-exchange market, for any of a variety of reasons: the limited number of banks involved, the brief time span of suspension, and especially the development of a more-integrated foreign-exchange market.[54] The 1857, 1873, 1893, and 1907 suspensions were of this nature. In the 1873, 1893, and 1907 periods, only the "deposit dollar"—not the "currency dollar," on which the exchange rate is based—noticeably depreciated against gold and foreign exchange. To the extent that the 1917–1922 and 1933–present paper standards involved fixed or managed exchange rates, exchange-rate variation was restricted.

Besides currency inconvertibility, divergences from the conditions for an international gold standard occurred during and around the paper standard occasioned by World War I, as shown in Table 17.3, and also with adoption of the 1933 paper standard. From August to October 1914, there was an informal embargo on the export of gold on the part of New York banks (Brown 1932, pp. 201–6). On September 7, 1917, by Presidential Proclamation, an embargo effective September 10 was imposed on exports of coin, bullion, and currency. Supported by the Treasury, Federal Reserve Board, and Postmaster General, the embargo lasted until June 9, 1919, when the Federal Reserve Board announced that licenses for the export of gold and currency would be freely granted (with rare geographical exceptions).[55] In 1933, on March 6, President Roosevelt suspended gold exports and on April 20 prohibited them (except by license).[56]

Complete exchange control was imposed under the paper standard of World War I: from January 1918 to June 1919, all foreign-exchange transactions required approval from the Federal Reserve Board (Taus 1943, pp. 154–5). Prohibition of bank dealings in foreign exchange was temporarily imposed under the 1933 paper standard.

Impounding of gold was a third common feature of the two paper standards. Through a combination of moral suasion and creative legislation over 1916–1918, gold and gold certificates were given up by commercial banks and concentrated in the Federal Reserve banks.[57] President

Table 17.3 Deviations from International Gold standard, 1914–1925

Deviation	*Time Period*	
	Britain	*United States*
Currency inconvertibility	August 1914–April 27, 1925[a]	April 6, 1917–March 17, 1922
Payments moratorium	August 2,1914–November 3, 1914	–
Restriction of gold exports	April 1, 1919–September 11, 1919[b]	August 2, 1914–October 1, 1914 September 10, 1917–June 8, 1919
Prohibition of gold imports[c]	December 5, 1916–March 31, 1919	–
Commandeering of gold imports	August 1914–September 11, 1919	–
Prohibition of gold melting	December 5, 1916–March 31, 1919[d]	–
Prohibition of buying or selling gold at a premium	May 18, 1918–December 30, 1925	–
Exchange control	December 24, 1914–January 14, 1924 November 1924–November 3, 1925	January 26, 1918–June 24, 1919
Exchange-rate management	August 1915– January 12, 1916	–
Exchange-rate pegging	January 13, 1916–March 19, 1919	–

[a]Currency convertibility restored on April 28, 1925, only for export of gold, not for domestic circulation
[b]Prohibition of export of gold produced outside British Empire continued to April 27, 1925
[c]Gold sold to Bank of England excluded
[d]Prohibition for non-Empire gold continued to April 27, 1925
Sources Atkin (1970, pp. 325–31), Beckhart (1924, pp. 267–8, 272–3), Brown (1929, pp. 6, 7, 20, 31, 37–41,47–8, 227–9; 1932, pp. 201, 204–6, 241–3, 248; 1940, pp. 31, 35, 37, 60, 180, 184, 378), Bullock, Williams, and Tucker (1919, p. 242), Cross (1923, pp. 377–81), Fraser (1933, pp. 33,40–1,45–6), Jaeger (1922, p. 24), Keynes (1930, vol. I, p. 19). Kirkaldy (1921, pp. 6–9, 33–4,421). Morgan (1952, pp. 12–13, 23, 64, 197–8, 261–5), Sayers (1976, pp. 55–6, 80–2). Spalding (1922, pp. 176–7). Taus (1943, p. 155)

Roosevelt's Proclamation of April 5, 1933 required all bank and non-bank owners of gold coin, gold bullion, or gold certificates to deliver all their present and future holdings, with minor exemptions, to a Federal Reserve bank, either directly or through commercial banks (member banks of the

Federal Reserve system), to be paid for at par.[58] A subsequent proclamation, of August 28, 1933, forbade anyone other than a Federal Reserve Bank from acquiring or holding gold in the United States or exporting gold, except under license. On December 28, 1933, the Secretary of the Treasury ordered that all gold be delivered to it at the official price of $20.671835 per fine ounce, resulting in almost a 70% profit with the $35 price instituted January 31, 1934. Throughout the 1934–1971 "limited gold-bullion standard," the holding of gold was forbidden to US residents, with minor exceptions.

In sum, the final two paper standards of the United States differed from all previous such episodes in that "suspension of specie payments," historically the only deviation from a metallic standard, was supported by a variety of other divergences from the international gold standard.

Notes

1. Later, the Reich (Holy Roman Empire) coined the "Reichsthaler," in English "rix-dollar," which in size and silver content was close to the peso. Therefore it was natural for the latter currency to be called the "Spanish dollar," originally by the London dealers in foreign exchange. See Nussbaum (1957, p. 10) and Carothers (1930, pp. 21–2).
2. The real was termed a "bit" by the colonists, a name they used for any small silver coin. The Spanish dollar was also divided into half-dollars and quarters, the latter known as "two bits." See Carothers (1930, pp. 34–5). Subsequently, the American (and Canadian) quarter-dollars were also colloquially called "two bits," a usage extending into the second half of the twentieth century.
3. The Morris report is in *International Monetary Conference* (1879, pp. 425–32) and is discussed by Carothers (1930, pp. 46–9) and Taxay (1966, pp. 15–16).
4. The Jefferson plan is reprinted in *International Monetary Conference* (1879, pp. 437–43) and discussed by Carothers (1930, pp. 50–1) and Taxay (1966, pp. 20–1).
5. The Congressional report and legislation, including a mint bill that followed in October 1786, are in *International Monetary Conference* (1879, pp. 445–51). For commentary, see Carothers (1930, pp. 51–6), Taxay (1966, pp. 22–5), and Watson (1899, pp. 19–25).
6. Hamilton's report is printed in *International Monetary Conference* (1879, pp. 454–84) and discussed by Carothers (1930, pp. 62–5) and Taxay (1966, pp. 44–51).

7. The Act is in Huntington and Mawhinney (1910, pp. 474–9), with good summaries provided by Carothers (1930, pp. 62–5) and Taxay (1966, pp. 65–7).
8. See Stewart (1924, pp. 18–19) and Carothers (1930, pp. 81–2). The individual states took their time in formally adopting the new federal unit of account—for example, Massachusetts in 1794, New York in 1797, Maryland in 1812, and. incredibly, New Hampshire not until 1948 and by popular referendum changing its constitution that established the shilling as the monetary unit. See Nussbaum (1957, p. 56).
9. On this differential treatment of gold and silver coins, see Alexander Hamilton's Report, in *International Monetary Conference* (1879, p. 456).
10. These comments are not to be interpreted as strictures against the colonial monetary system. Indeed, as Sylla (1982, p. 23) writes: "One would be hard pressed to find a place and time in which there was more monetary innovation than in the British North American colonies in the century and a half before the American Revolution." To the accomplishments mentioned by Sylla one might add, for the Revolutionary Period, the joint Congressional-state refunding plan of 1780 for the redemption of Continental bills. This was historically the first contractionary monetary reform of a paper currency. For the basic literature on the colonial monetary system, see the bibliographical essay prepared by Perkins (1980, pp. 121–2). The paper-money experience of the Revolutionary Period is described by Carothers (1930, pp. 37–41), Dewey (1934, pp. 34–41), Nussbaum (1957, pp. 35–9), Studenski and Krooss (1963, pp. 25–9), Nettles (1962, pp. 24–31), Hepburn (1924, pp. 13–19), Sumner (1874, pp. 43–9), and Del Mar (1899, pp. 93–116). As Sylla points out, it was the reaction to the inflationary paper-money experience of the Revolution (rather than the mixed history of colonial paper money) that led to a specie standard for the federal United States. The Continental bills ("old tenor") depreciated to one-thousandth of face value, becoming worthless by 1780. Sumner (1874, pp. 46–7) writes: "A barber's shop in Philadelphia was papered with it, and a dog, coated with tar, and the bills stuck all over him, was paraded in the streets." Even the reform currency ("new tenor") depreciated to one-sixth of its silver value.
11. Whether under troy weight, used for precious metals, or the common avoirdupois weight, a grain is identical. Under troy weight, there are 24 grains in a pennyweight, 480 in an ounce, and 5760 in a pound. In contrast, there are 437.5 grains in an avoirdupois ounce and 7000 in a pound.
12. The reports are reprinted and discussed in Watson (1899, pp. 21–2, 243–55).
13. Hamilton's report is impressive in its basis on economic argument. One can agree with Watson (1899, p. 33) that: "It is difficult to pay to this

report that tribute which it deserves. It was so exhaustive in its analysis, so profound in its reasoning, so comprehensive and logical in every position taken, that to this day it is regarded as an authority on money and coinage."

14. For the explanation of why the unwieldy fineness was adopted, see Carothers (1930, pp. 62–3), Willem (1959, pp. 4–6), and Kemmerer (1944, p. 66, n. 3). The legislations are in Huntington and Mawhinney (1910, pp. 473–9).
15. The date traditionally ascribed to this emission is October 1792 (see, for example, Watson, 1899, p. 64; Hepburn, 1924, p. 45); but Taxay (1966, pp. 71–2) provides evidence that it was no later than July.
16. The letter, dated December 30,1893, is printed in Lowrie and Clarke (1832, pp. 270–1). The legislation is in Peters (1848, p. 341). For a description of this episode, see Taxay (1966, pp. 120–1). Until 1873 the Mint was attached to the Department of State rather than the Treasury.
17. The first gold deposit occurred on February 12, 1795, with its coinage done July 31. For a list and description of the early deposits at the Mint, see Stewart (1924, pp. 44–50).
18. The Mint officials involved in the over-fineness of the silver dollar were not punished—not even reprimanded—for their behavior. Depositors, of course, received less coined money for their silver bullion than the law specified. One such depositor received reimbursement from Congress. Discussions of the over-fine silver-dollar episode are presented in Bolles (1894, vol. II, pp. 161–3), Taxay (1966, pp. 89–90), Watson (1899, pp. 229–31), and Willem (1959, pp. 1–9). Strangely, Taxay and Willem state that the 1794–1795 dollar consisted of 374.75 grains of fine silver (rather than the true 374.4). The source documents—printed in Lowrie and Clarke (1832, pp. 352–8, 588), Congress of the United States (1851, pp. 3667–71), and Select Committee on Coins (1832, pp. 17–20)—clearly show 374.4 to be the correct number. Hepburn (1924, p. 44, n. 1) explicitly states the correct figure, while Watson quotes a source document containing it.
19. The Acts are in Huntington and Mawhinney (1910, pp. 496–7, 500–8) and discussed in Carothers (1930, pp. 91–5), Taxay (1966, p. 200), and Watson (1899, pp. 85–7, 97–9).
20. The acts mentioned in this paragraph are in Huntington and Mawhinney (1910, pp. 508–9, 511–13, 530–50, 593, 610–14). Time series of gold and silver coinage are in Director of the Mint (1942, pp. 68–71). On the one-dollar piece, see Carothers (1930, p. 135).
21. The Legislation and Proclamation are in Krooss (1969, vol. iv, pp. 2793–805).

22. For the text of some of these Acts, see Huntington and Mawhinney (1910, pp. 481–91, 497–8). A summary and critique are provided by Carothers (1930, pp. 66–7, 78–9, 101–2). See also Hepburn (1924, pp. 46–7, 60).
23. For the act, see Huntington and Mawhinney (1910, pp. 517–18); for the history of withdrawing the foreign coin, Carothers (1930, pp. 138–48).
24. The Act is in Huntington and Mawhinney (1910, pp. 511–13).
25. See Huntington and Mawhinney (1910, p. 568) and Laughlin (1900, p. 305).
26. See, for example, Carothers (1930, p. 233) and Friedman (1990b, p. 1165).
27. "According to bimetallists, the Coinage Act of 1873, which discontinued the silver dollar as a monetary standard, passed Congress through the corrupt influence of a cabal of powerful government bondholders who conspired with treasury officials and influential congressmen. By establishing a single gold unit of account, the cabal presumably hoped to raise the market value of its public securities" (Weinstein, 1967, p. 307). Histories of "silver politics" are provided by Friedman and Schwartz (1963, pp. 113–19), Hepburn (1924, pp. 268–304), Laughlin (1900, pp. 92–105, 211–17, 259–61), and Myers (1970, pp. 197–222).
28. The silver legislations are in Huntington and Mawhinney (1910, pp. 689–90, 579–81, 589–91, 599–600).
29. For all the above legislation, see Huntington and Mawhinney (1910, pp. 634–45) and Sanger (1863, p. 338).
30. There is evidence that certificate issuance had always been of this nature. See Friedman and Schwartz (1963, p. 25, n. 11).
31. These Acts pertaining to gold certificates are in Huntington and Mawhinney (1910, pp. 175–9, 693–6, 704–10) and *Statutes at Large of the United States of America* (1911, p. 965; 1921, p. 370).
32. The Federal Reserve Act and the Act of June 5, 1933 Are in Krooss (1969, vol. iv, pp. 2436–70, 2723–4).
33. Morris' money unit was 0.25 grain of silver; therefore one pound (5760 grains) of silver would coin 23,040 units. The mint price, however, would be 22,237 units per pound. Mint charges, therefore, would be (23,040 − 22,237)/23,040 = 3.49%. Under the 1786 Acts, a silver (gold) dollar would have a zero-charge mint price of 5760/409.7891 (5760/26.8656) dollars per pound Troy, equaling \$14.0560 (\$214.4006), versus an actual mint price of \$13.777 (\$209.77), involving mint charges of (14.0560 − 13.777)/14.0560 = 1.98% for silver and (214.4006 − 209.77)/214.4006 = 2.16% for gold.
34. The logical interpretation, given by Bolles (1894, vol. II, pp. 512–13), that payment in coin is to be provided within 40 days of the deposit of bullion and within five days for the 0.5% charge, is not the letter of the statute.

35. However, the Act of March 31, 1795 did assess charges by weight for deposits of bullion below US standard. See Huntington and Mawhinney (1910, pp. 483–5).
36. Coinage itself actually occurred at the Mint, not at the Assay Office. All other Mint functions were performed on the premises of the Assay Office. For discussion of the Assay Office at New York, see Watson (1926, pp. 10–12, 19, 32–3).
37. For these Acts, See Huntington and Mawhinney (1910, pp. 586, 596, 616).
38. For the early history of banking in the United States, see Hammond (1957, pp. 40–88). The statistics are from Hammond (1957, pp. 144–6).
39. See Huntington and Mawhinney (1910, pp. 327–9, 330–62) for these Acts, and Friedman and Schwartz (1963, pp. 16–23) for discussion of the formative years of the national banking system.
40. The Bland-Allison Act of February 28, 1878 authorized issuance of silver certificates by the Treasury in return for deposits of silver dollars.
41. See Nussbaum (1950, pp. 596–7). The Act of June 12, 1945 is in Krooss (1969, vol. iv, pp. 2875–6).
42. There are two cases in which legal bimetallism can be effective. First, if the country happens to select a mint ratio close to the world gold/silver price ratio, bimetallism results as long as the divergence of the ratios is within limits (set by arbitrage costs and market imperfections). Friedman (1990a, p. 90) notes that "these costs define upper and lower 'gold-silver price ratio points' between which the market ratio can vary without producing the complete replacement of one metal by the other [in the domestic country]." Such a situation could likely exist only temporarily. England had effective bimetallism for a few years at the turn of the eighteenth century, in the process of switching from an effective silver to an effective gold standard (see Officer 1996, Sect. 8 of chapter 4).

    Second, a lasting bimetallism can happen if the country possesses a sufficient stock of gold and silver coin and is important enough in the international economy to dominate the world gold/silver price ratio. France, with a mint ratio of 15.5, was in a position of dominance from 1803 to 1850 (Yeager, 1976, p. 296). Friedman (1990a, p. 89) points out that what gave France preponderant influence on the world price ratio (although he exaggerates in describing France's ability to "peg" the ratio) was not only France's economic importance relative to the rest of the world but also the country's high propensity to use specie as money, both directly as coins and indirectly as reserves for paper currency and bank deposits.
43. The market series, compiled by Soetbeer (1879, pp. 130–1), is of much higher quality than alternative data. The Soetbeer series is an annual average of twice-weekly official market quotations in Hamburg to 1832

and uses generally accepted London data thereafter. In contrast, alternative series (based on the London market) exhibit neither their data source nor their method of construction and furthermore suffer from obvious errors, both in the level of some observations and in their year-to-year movement. See S. Dana Horton, various appendices, in *International Monetary Conference* (1879, pp. 649, 701, 708–9) and Laughlin (1896, pp. 288–91). The French mint ratio of 15.5 is not used in place of the Soetbeer data because, while the world price ratio may have been principally determined by the French ratio, deviations did occur and in fact were the norm. Indeed, for a minority view claiming that the reach of French bimetallism has been exaggerated, see Shaw (1896, pp. 178–80).

44. His average computed market ratio in the United States was 14.99, the world ratio in 1791 was 15.05, and the recommended and adopted ratio 15.
45. Again the Soetbeer data are used for the 1834–1873 market rate. The two episodes (1792–1834 and 1834–1873) of the divergence between legal and market rates are described by Carothers (1930, pp. 75, 81–101), Hepburn (1924, pp. 47–61), and Watson (1899, pp. 71–3, 78–96).
46. Actually, termination of coinage of the silver dollar in 1873 and its demonetization in 1874 merely reflected long-standing reality. As Carothers (1930, p. 235) states: "Exported before 1806, not coined from 1806 to 1836, and not in circulation from 1836 to 1873, the [silver] dollar was an unknown coin." (On this history of the silver dollar, see Officer 1996, note 10 to chapter 5.) So the elimination of the silver dollar by the 1873–1874 legislation was in the nature of modifying the coinage law to accord with actuality. However, it is argued by Friedman (1990b) that the economic consequences of abandoning legal bimetallism, which meant alternating effective monometallism, were harmful for the United States. In particular, an effective silver standard would have avoided the 1891–1897 crisis (see Officer 1996, Sect. 2 of chapter 13).
47. For discussion of the meaning and mechanism of suspension of specie payments, see Temin (1969, pp. 114–18) and Triffin (1960, p. 22). This phenomenon was also part of the colonial experience with paper money. The colonies did not set fixed rates between paper currency and coin, and markets existed in which the two types of money traded for each other. Perkins (1980, p. 111) writes of "the market value of the paper relative to specie and foreign exchange" in the colonial period. He notes that in situations in which paper had depreciated, creditors would accept either specie or paper money at its current market value. In contrast, during the Continental-money experience of the Revolutionary Period, Gresham's Law operated in full force and coin disappeared from circulation. The reason is that the states legislated strict parity of the paper with coined money. The penalties for not respecting the face value of

Continental currency were severe. “The notes were made full legal tender and refusal to accept them forfeited the debt and incurred other money penalties, pillory. imprisonment, loss of ears even, and being outlawed as enemies of their country” (Hepburn, 1924, p. 17).

48. “Each of the antebellum financial panics had been marked by a rush on diverse state-chartered banks by holders of banknotes who wanted to convert their holdings into specie…after the Civil War the public did not exhibit doubts about the safety of national banknotes. The postbellum panics were instead marked by rushes of bank depositors to convert their deposits into currency, which included national banknotes and federal government issues as well as specie” (Sylla 1972, p. 233).
49. For histories of these suspensions, see (1) *for 1814–1817*, Bolles (1894, vol. II, pp. 261–83, 317–29), Hammond (1957, pp. 227–50), Secretary of the Treasury (1838, p. 5), Smith (1953, pp. 110–15), Smith and Cole (1935, pp. 25–9), and Sumner (1874, pp. 64–75); (2) *for 1837–1842*, Davis and Hughes (1960, pp. 57, 61), Hammond (1957, pp. 465–501),Hepburn (1924, pp. 132–8), Knox (1900, pp. 76–7, 502–5), Martin (1898, pp. 30–3), Myers (1931, pp. 64–8), Smith (1953, pp. 190–227), Sumner (1874, pp. 132–54), and Temin (1969, pp. 113–71); (3) *for 1857*, Dunbar (1904, pp. 266–93), Hammond (1957, pp. 710–13), Knox (1900, pp. 512–13), and Sumner (1874, pp. 180–7); (4) *for 1860–1861*, Dunbar(1904, pp. 309–10) and Knox (1900, pp. 513–14); (5) for *1861–1878*, Officer (1981) and the references cited there.
50. On June 17, 1864 Congress legislated the prohibition of the gold market (though allowing brokers to transact in gold within their offices), but the bill was repealed on July 2 (Huntington and Mawhinney, 1910, pp. 182–3). The intent was to reduce the premium on gold, but the law was ineffective, with the premium actually increasing.
51. On the 1873, 1893, and 1907 currency premiums, see Andrew (1908, pp. 290–3), Clark (1984, pp. 819–20), Cross (1923, pp. 397–9), Friedman and Schwartz (1963, pp. 110, 161–2), and Sprague (1910, pp. 56–61,186–95, 280–6).
52. See Chandler (1958, pp. 103–4), Secretary of the Treasury (1920, pp. 181–2; 1922, p. 72), Brown (1940, p. 37), and Beckhart (1924, p. 267).
53. Excellent histories of the 1933–1934 events are provided by Friedman and Schwartz (1963, pp. 462–74) and Yeager (1976, pp. 346–50).
54. The last explanation is stated by Davis and Hughes (1960, p. 62) and Perkins (1975, pp. 155–6).
55. See Taus (1943, p. 153), Cross (1923, p. 377), Brown (1929, pp. 18, 26; 1940, pp. 34, 37), and Beckhart (1924, pp. 268–73).

56. The latter was by Proclamation under the Act of October 6, 1917, as amended by the Act of March 9, 1933. The Proclamation is in Krooss (1969, vol. 4, pp. 2717–18).
57. For details, see Beckhart (1924, pp. 252–67),Taus (1943, pp. 157, 178), Chandler (1958, pp. 102–3), Brown (1940, p. 43), and Cross (1923, p. 377).
58. The Proclamation is in Krooss (1969, vol. 4, pp. 2714–16).

References

Andrew, A. Piatt. "Hoarding in the Panic of 1907." *Quarterly Journal of Economics* 22 (February 1908), pp. 290–9.

Atkin, John. "Official Regulation of British Overseas Investment, 1914–1931." *Economic History Review* 23, second series (August 1970), pp. 324–35.

Beckhart, Benjamin Haggott. *The Discount Policy of the Federal Reserve System*. New York: Henry Holt and Company, 1924.

Bolles Albert S. *The Financial History of the United States*, vol. II, From 1789 to 1860, 4th edn; vol. III, From 1861 to 1885, 2nd edn. New York; D. Appleton & Company, 1894.

Brown, William Adams, Jr. *England and the New Gold Standard 1919–1926*. London: P. S. King & Son, 1929.

______. "The Government and the Money Market." In Benjamin Haggott Beckhart, James G. Smith, and William Adams Brown, Jr. *The New York Money Market*. vol. IV, External and Internal Relations, part 2. New York: Columbia University Press, 1932.

______. *The International Gold Standard Reinterpreted*, 2 vols. New York: National Bureau of Economic Research, 1940.

Bullock, Charles J., John H. Williams, and Rufus S. Tucker, "The Balance of Trade of the United States." *Review of Economic Statistics* 1 (July 1919), pp. 215–66.

Calmoris, Charles W. and Larry Schweikart. "The Panic of 1857: Origins, Transmission, and Containment." *Journal of Economic History* 51 (December 1991), pp. 807–34.

Carothers, Neil. *Fractional Money*. New York: John Wiley & Sons, 1930.

Chandler, Lester V. *Benjamin Strong, Central Banker*. Washington, DC: The Brookings Institution, 1958.

Clark, Truman A. "Violations of the Gold Points, 1890–1908." *Journal of Political Economy* 92 (October 1984), pp. 791–823.

Committee on Commerce. "Branch Mint, New York [Phoenix Report]." House Report No. 490, 31st Cong.,1st session, September 14, 1850.

Congress of the United States. *Annals*, 5th Cong. Washington, DC: Gales and Seaton, 1851.

Cross, Ira B. *Domestic and Foreign Exchange*. New York: The Macmillan Company, 1923.

Davis, Lance E. and Jonathan R. T. Hughes. "A Dollar-Sterling Exchange, 1803–1895." *Economic History Review* 13 (August 1960), pp. 52–78.

Del Mar, Alexander. *The History of Money in America from the Earliest Times to the Establishment of the Constitution*. New York: The Cambridge Encyclopedia Company, 1899.

de Vries, Margaret Garritsen. *The International Monetary Fund 1972–1978*. vol. I, Narrative and Analysis. Washington, DC: International Monetary Fund, 1985.

Dewey, Davis Rich. *Financial History of the United States*, 12th edn. New York: Longmans, Green and Co., 1934.

Director of the Mint. *Annual Report*. Washington, DC: Government Printing Office, 1895, 1942.

Dunbar, Charles Franklin. *Economic Essays*. New York: The Macmillan Company, 1904.

Fraser, Herbert F. *Great Britain and the Gold Standard*. London: Macmillan and Co., 1933.

Friedman, Milton. "Bimetallism Revisited." *Journal of Economic Perspectives* 4 (Fall 1990a), pp. 85–104.

______. "The Crime of 1873." *Journal of Political Economy* 98 (December 1990b), pp. 1159-94.

Friedman, Milton, and Anna Jacobson Schwartz. *A Monetary History of the United States 1867–1960*. Princeton, NJ: National Bureau of Economic Research, Princeton University Press, 1963.

Hammond, Bray. *Banks and Politics in America*. Princeton, NJ: Princeton University Press, 1957.

Hepburn, A. Barton. *A History of Currency in the United States*. New York: The Macmillan Company, 1924.

Huntington, Andrew T., and Mawhinney, Robert J., eds. *Laws of the United States Concerning Money, Banking, and Loans, 1778–1909*. Senate Document No 580, 61st Cong., 2nd session. Washington, DC: Government Printing Office, 1910.

International Monetary Conference. Senate Executive Document No. 58, 45th Cong. 3rd session. Washington, DC: Government Printing Office 1879.

Jaeger, Ruth M. *Stabilization of the Foreign Exchanges*. Ph.D. dissertation. New York: Columbia University, 1922.

Kemmerer, Edwin Walter. *Gold and the Gold Standard*. New York: McGraw-Hill Book Company, 1944.

Keynes, John Maynard. *A Treatise on Money*, 2 vols. London: Macmillan and Co., 1930.

Kirkaldy, Adam W. *British Finance During and After the War, 1914-21.* London: Sir Isaac Pitman & Sons, 1921.

Knox, John Jay. *A History of Banking in the United States*. New York: Bradford Rhodes, 1900.

Krooss, Herman E., ed. *Documentary History of Banking and Currency in the United States*, 4 vols. New York: Chelsea House, 1969.

Laughlin, J. Laurence. *The History of Bimetallism in the United States*. New York: D. Appleton and Company, 1896; 4th edn., 1900.

Lowrie, Walter, and Matthew St. Clair Clarke. *American State Papers*, vol. V. Washington: Gales and Seaton, 1832,

Martin, David A. "1853: The End of Bimetallism in the United State," *Journal of Economic History* 33 (December 1973), pp. 825–44.

Martin, Joseph G. *A Century of Finance. Martin's History of the Boston Stock and Money Markets, One Hundred Years*. Boston: privately printed, 1898.

McGrane, Reginald Charles. *The Panic of 1837*. New York: Russel & Russel, 1965.

Merchants' Magazine 1 (July 1839), 19 (September 1848).

Mitchell, Wesley Clair. *A History of the Greenbacks*. Chicago: University of Chicago Press, 1903.

Morgan, E. Victor. *Studies in British Financial Policy, 1914–25*. London: Macmillan & Co., 1952,

Myers, Margaret G. *The New York Money Market*, vol. I, Origins and Development. New York: Columbia University Press, 1931.

______. *A Financial History of the United States*. New York: Columbia University Press, 1970.

Nettles, Curtis Putnam. *The Emergence of a National Economy*. New York: Holt, Rinehart and Winston, 1962.

Nussbaum, Arthur. *Money in the Law: National and International*. Brooklyn, NY: The Foundation Press, 1950.

______. *A History of the Dollar*. New York: Columbia University Press, 1957.

Officer, Lawrence H. "The Floating Dollar in the Greenback Period: A Test of Theories of Exchange-Rate Determination." *Journal of Economic History* 41 (September 1981), pp. 629–50.

______. *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*. Cambridge: Cambridge University Press, 1996.

Perkins, Edwin J. *Financing Anglo-American Trade*. Cambridge, MA: Harvard University Press, 1975.

______. *The Economy of Colonial America*. New York: Columbia University Press, 1980.

Peters, Richard, ed. *Public Statutes at Large of the United States of America*. Boston: Charles C. Little and James Brown, vols. 1–2, 1848; vols. 3–4, 1846. Little Brown and Company, vols. 4–5, 1860.

Sanger, George P., ed. *Public Laws of the United States of America*, 1862–1863, 1863–1864, 1864–1865, 1869–1870. Boston: Little, Brown and Company, 1863, 1864, 1865, 1870.

Sayers, Richard S. *The Bank of England 1891–1914*. Appendixes. Cambridge: Cambridge University Press, 1976.

Secretary of the Treasury. *Report Respecting the Relative Value of Gold and Silver, etc*. House Document No. 117, 21st Cong., 1st session, May 29,1830.

______. *Report...in Compliance with a Resolution of the Senate Relative to the Means of Supplying the Mint with Bullion and Foreign Coins for Coinage*. Senate Document No. 162, 24th Cong., 1st session, February 15, 1836.

______. *Report...Transmitting Statements of the Rates of Exchange and Prices of Bank Notes at Different Periods*. Senate Document No. 457, 25th Cong., 2nd session, May 28, 1838.

______. *Annual Report*, 1900, 1920, 1922, 1926. Washington, DC: Government Printing Office, 1901, 1921, 1923, 1927.

Select Committee on Coins. "Gold and Silver Coins." House of Representatives, Report No. 496, 22nd Cong., 1st session. In *Reports of Committees of the House of Representatives*, vol. V. Washington, DC: Duff Green, 1832.

Shaw, William A. *The History of Currency 1252 to 1894*, 2nd edn. New York: G. P. Putnam's Sons, 1896.

Smith, Walter Buckingham. *Economic Aspects of the Second Bank of the United States*. Cambridge, MA: Harvard University Press, 1953.

Smith, Walter Buckingham, and Arthur Harrison Cole. *Fluctuations in American Business, 1790–1860*. Cambridge, MA: Harvard University Press, 1935.

Soetbeer, Adolf. *Edelmetall-Produktion*. Gotha: Justus Perthes, 1879.

Spalding, William F. *The London Money Market*, 2nd edn. London: Sir Isaac Pitman & Sons, 1922.

Sprague, Oliver M. W. *History of Crises under the National Banking System*. Senate Document No. 538, 61st Cong., 2nd session, Washington, DC: Government Printing Office, 1910.

Statutes at Large of the United States of America. August 1893 to March 1895, vol. 28; March 1909 to March 1911, vol. 36, part 1; May 1919 to March 1921, vol. 41, part 1; April 1921 to March 1923, vol. 42, part 1. Washington, DC: Government Printing Office, 1895, 1911, 1921, 1923.

Stewart, Frank H. *History of the First United States Mint*. Camden, NJ: privately printed, 1924.

Studenski, Paul, and Herman E. Krooss. *Financial History of the United States*, 2nd edn. New York: McGraw-Hill Book Company, 1963.

Sumner, William Graham. *A History of American Currency*. New York: Henry Holt and Company, 1874.

Sylla, Richard. "The United States 1863-1913." In Rondo Cameron, ed. *Banking and Economic Development*. New York: Oxford University Press, 1972, pp. 232-62.

______. "Monetary Innovation in America." *Journal of Economic History* 42 (March 1982), pp. 21–30.

Taus, Esther Rogoff. *Central Banking Functions of the United States Treasury, 1789–1941.* New York: Columbia University Press, 1943.

Taxay, Don. *The US Mint and Coinage*. New York: Arco Publishing Company, 1966.

Temin, Peter. *The Jacksonian Economy*. New York: W. W. Norton & Company, 1969.

Triffin, Robert. *Gold and the Dollar Crisis*. New Haven: Yale University Press, 1960.

Watson, David K. *History of American Coinage*. New York: G. P. Putnam's Sons, 1899.

Watson, Jesse P. *The Bureau of the Mint*. Baltimore: The Johns Hopkins Press, 1926.

Weinstein, Allen. "Was There a 'Crime of 1873'?: The Case of the Demonetized Dollar." *Journal of American History* 54 (September 1967), pp. 307-26.

Willem. John M„ Jr. *The United States Trade Dollar*. New York: privately printed, 1959.

Yeager, Leland B. *International Monetary Relations: Theory, History, and Policy*. New York: Harper & Row, 1976.

CHAPTER 18

British Monetary Standard

Originally published in *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*, pp. 34–46, 291–293, copyright Cambridge University Press, 1996. Reprinted with permission.

18.1 Unit of Account

The pound, shilling, and pence schemata as the unit of account began prior to the Norman Conquest, in the Anglo-Saxon period, when the only coin in existence was the silver penny.[1] With rare and temporary exception, the penny was the only coin in circulation until the late thirteenth century. The term "penny" is believed to emanate from a coin issued by King Penda of Mercia in the seventh country. In the ninth century the penny was in all the Saxon kingdoms and 240 pennies always constituted a pound. Even though the weight of the penny varied across kingdoms and over time, "the intention was that the pound weight of silver and the pound of money should be the same and that the pound of silver should be minted into 240 pennies" (Feavearyear 1963, pp. 7–8). The abbreviation for penny, d., is from the Latin *denarius* (a Roman coin and unit of account), which until the fourteenth century meant both a pennyweight and the face value of the penny coin.

According to Feavearyear, the Saxons may have applied the term "scilling," later "shilling" and meaning "a piece cut off," to pieces of broken silver added to the scale to compensate for underweight coin.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_18

Over time, the shilling became equal to a definite number of pennies, eventually 12, the preference of William the Conqueror. Following Roman nomenclature, the Normans determined the unit of account as the *libra* (the ancient Roman pound, weight measure, abbreviated as £), equal to 20 *solidi* (a Roman coin, abbreviated as s.), with the *solidus* equal to 12 *denarii*. In English: pound, shilling, and pence.[2]

18.2 Mint Price and Coins Struck

In sharp contrast to the American experience, in which a national mint was not functioning until 1794, the minting of coined money had been undertaken by British monetary authorities more or less continuously since Roman times.[3] The penny declined in size steadily over time from 24 grains in the eighth century to a third that weight in 1601. Notwithstanding its role in the unit of account, the shilling as a coin was not issued until 1504; a variety of other silver coins before and after were also produced (Feavearyear 1963, p. 439). With rare exception, (1) all silver coins were made of "sterling silver," that is, they were of 37/40th (or 0.925) fineness, and (2) the value/weight ratio of the other coins was proportional to that of the penny. Coining of silver was prohibited by the Act of June 21, 1798, but restored by the Act of June 22, 1816 (known as Lord Liverpool's Coinage Act).[4]

Gold coins were first issued in 1257, and more than a dozen different coins subject to varying weight and denomination were produced until the two final and most important: the guinea and sovereign, which entered circulation in 1663 and 1817, respectively (Feavearyear 1963, pp. 437–438).[5] The guinea had initial value and the sovereign continuous value of 20 shillings (as did the old sovereign from 1489 to 1526, the first coin with such valuation), thus representing the pound. Other denominations of these two basic coins were also produced. The fineness of gold coins varied with the specific coin until the guinea, when an 11/12th (or 0.9166…) fineness was established and retained for the sovereign.

In 1560 silver was coined at 60d. per standard ounce (stated as: sterling silver coined at 60d. per ounce). Two gold coins, with different valuations, were in existence: the angel and crown, coined at 723.7d. and 720d., respectively, per fine ounce. The mint (gold/silver price) ratio was the gold valuation [723.7d. or 720d.] divided by the silver valuation converted to a fine basis [(60d.) · (40/37)]: 11.16 for the angel and

11.10 for the crown. On September 29, 1601, both metals were increased in value at the mint (all coins reduced in weight), silver in particular to 62d. per standard ounce and gold proportionately less, so the mint ratios fell to 10.95 and 10.90. Further increases in the value of gold, holding silver constant, then increased the mint ratio to 14.25 and 14.21 for the angel and crown, respectively, by 1662 (Horsefield 1960, p. 75).

Considered another way, the mint indenture of 1601 involved one pound Troy, or 5760 grains, of standard silver coined into 62 shillings, whence the silver weight of the pound sterling was (20/62) · (5760) · (37/40) = 1718.7096+ (1718 22/31) fine (pure) grains. On December 24, 1663, King Charles II ordered the guinea to be coined.[6] By the Order, one pound Troy, or 5760 grains, of standard gold was made into 44.5 guineas, implying a guinea weight of 129.4382+ (129 39/89) grains, and the value of the guinea was set at 20 shillings. At this valuation, the gold weight of the pound sterling was (129.4382+) · (11/12) = 118.6516+ (118 58/89) fine grains. The silver-weight to gold-weight ratio is the mint ratio, which was 14.49.

Only government offices had to respect the official rating of the guinea (or of any gold coin); private transactions took place at market rates. On March 7, 1696, in an effort to reduce the market price, an Act was passed forbidding any person from transacting in guineas at a rate higher than 26 shillings per guinea, effective March 25, 1696. The gold weight of the £ became (129.4382+) · (11/12) · (20/26) = 91.2705+ (91 313/1157) fine grains, and the mint ratio increased to 18.83. A subsequent Act lowered the price ceiling to 22 shillings from April 10, 1696.[7] This resulted in a £ gold weight of (129.4382+) · (11/12) · (20/22) = 107.8651+ (107 77/89) fine grains and a mint ratio of l5.93.

On February 15, 1699, the Treasury ordered that, in the receipt of taxes, guineas be accepted at no more than 21s. 6d., the day before Parliament agreed to a report to that effect.[8] The £ gold weight became (129.4382+) · (11/12) · (20/21.5) = 110.3736+ (110 1430/3827) fine grains, and the mint ratio fell to 15.57. In response to a Parliamentary resolution of December 21, 1717, a Royal Proclamation was issued the next day lowering the maximum price of the guinea in transactions to 21 shillings.[9] This valuation established the gold content of the £ at (129.4382+) · (11/12) · (20/21) = 113.0016+ (113 1/623) fine grains and the resultant mint ratio at 15.21.

The Coinage Act of 1816 stated that the weight and fineness of gold coin should be that prescribed by the existing mint indenture (meaning that for the guinea). On July 1, 1817, a Royal Proclamation revealed the ordering of coinage of the sovereign, weighing 123.274 grains of standard gold, to have the valuation of 20 shillings.[10] The accurate weight of the sovereign consistent with the guinea is rather $(129.4382+) \cdot (20/21) = 123.274478+$ (123 171/623), which the Coinage Act of April 4, 1870 expressed truncated to five decimal places, as 123.27447.[11] That Act decreed that 934.5 sovereigns should be coined from 20 pounds, or (20) · (5760) grains, of standard gold. The gold weight of the pound is then $(1/934.5) \cdot (20) \cdot (5760) \cdot (11/12) = 113.0016+$ (113 1/623) grains of pure gold, the same as that established in 1717. In fact, the Act of 1870 specifically continues the mint indenture of the Coinage Act of 1816, which was that for the guinea.

The Coinage Act of 1816 did reduce the weight of silver coins, with one pound Troy coined into 66 rather than 62 shillings, but silver became a subsidiary coinage and the mint ratio ceased to have economic meaning. The gold content of the pound, like that of the dollar, remained unchanged through World War I and even after 1925, the last year in which gold was coined. With 123.2744+ (123 171/623) grains of standard gold constituting £1 and 480 grains to the ounce, the mint price of standard gold from December 22, 1717 onward was 480/(123.2744+) = £3.89375, or £3 17s. 10.5d. per ounce.

18.3 Legal-Tender Status of Coin

Until the late eighteenth century, the legal-tender quality of coin is found not in specifically oriented statutes but in Royal Proclamations, mint indentures, and incidental passages in legislation.[12] Phrases such as "current coins," "lawful coin," and "current money" established or accepted the legal-tender status of gold and silver coin produced by the Royal Mint. It follows that with the coinage of gold as well as silver, legal bimetallism resulted. However, because of the practice of expressing the legal denomination of gold coins in terms of shilling and pence, in a sense there was an underlying silver standard.

Formal bimetallism was disturbed by the Act of May 10, 1774, which temporarily limited the legal tender of silver coin to £25; above that amount it could be tendered only by weight and at a maximum valuation of 62 pence an ounce. The act was renewed but allowed to lapse on May 1, 1783, then renewed on June 21, 1798, and made perpetual on July 12, 1799.[13] The Coinage Act of 1816 restored the legal tender of silver coin but only up to £2 and eliminated its tender by weight. Gold coin was made the only standard of value and its legal tender was explicitly stated as unlimited. Britain was now legally on the gold standard. The Coinage Act of 1870 repeated the legal-tender provisions of the 1816 legislation. Ironically, the statute that placed Britain on the gold standard occurred during a paper-standard regime, the Bank Restriction Period (see Sect. 18.9 below), just as the formal American move to a gold standard happened during the greenback paper-standard period.

18.4 Convertibility of Government/Central-Bank Paper into Coin

Immediately after its founding in 1694, the Bank of England issued banknotes. Of course, like any debtor, the Bank was required to discharge this liability in legal tender, namely, gold or silver coin. It was only with the Resumption Act (Peel's Act) of July 2, 1819 that specific statutory content was given to the obligation of the Bank to redeem its notes in gold coin on demand. The Bank of England Act of August 29, 1833 made Bank of England notes for sums above £5 legal tender (beginning August 1, 1834), as long as the Bank continued to pay its notes in coin on demand. Finally, the Currency and Bank Notes Act of July 2, 1928 made the Bank's notes legal tender in any amount.

Government paper money was not authorized until the Currency and Bank Notes Act of August 6, 1914, which permitted the issuance of currency notes with full legal-tender status and cashable into gold coin on demand at the Bank of England. The 1928 Act declared that currency notes are deemed to be Bank of England notes, thus amalgamating the two paper currencies.[14]

18.5 Nature of Coinage

Traditionally, coinage at British mints was open to all, although from the late twelfth to the fourteenth century silver could be coined only at king's exchanges, which levied an additional fee. "An Act for encouragement of coinage," effective December 20, 1666, kept continuously in force by subsequent legislation (except for a temporary lapse in 1680–1685) and made perpetual in 1768, explicitly stated freedom of coinage for everyone: "whatsoever person or persons, native or foreigner, alien or stranger."[15] The Act of June 21, 1798, made perpetual by the Act of July 12, 1799, terminated the coinage of silver. Free coinage of silver was technically restored in the Coinage Act of 1816—but only subsequent to a Royal Proclamation to that effect, which was never issued—and was formally eliminated in the Coinage Act of 1870.

Free coinage of gold was undisturbed by the 1870 Act; but the minimum amount of bullion required was always much higher than in the United States, where, prior to the Mint Act of 1873, there was no minimum, and by that Act bullion of value less than $100 could be refused. The British minimum is stated as £10,000 for 1817–1820 (Craig 1953, p. 289) and £20,000 in the mid-nineteenth and the twentieth centuries (Seyd 1868, p. 158; Spalding 1928, p. 86).

In England, as in every other country, there was a charge for coinage, consisting of mint expenses plus seigniorage for the king. The charge was variable over time and, with the important exception of part of the sixteenth century, generally low in percentage terms.[16] The Act of 1666 abolished mint charges, specifically those for melting and coining, for both gold and silver. England thus became the first country to have free and gratuitous coinage. Further, whether the fineness of bullion was above or below standard, no charge was to be made and payment was to be altered proportionately. The assayer and porter received "small fees" (Craig 1953, p. 239). Also, the mint charged for melting into bars, if the bullion was not already in ingot form. The Act of 1870 repeated the absence of mint charges for gold bullion, but provided that the mint could refuse bullion that required refining to achieve standard fineness.

The Act of 1816 established that silver would be coined at 66 shillings per pound Troy but depositors would be paid the old rate of 62 shillings. Thus there would be a seigniorage of 6.06% for silver coinage. However, this provision was a dead letter, because the necessary proclamation to permit private depositing of silver at the mint was never issued.

Waiting time at the mint could be eliminated between the late twelfth and sixteenth centuries by paying an extra charge at the king's exchanges. In the fourteenth century the delay between deposit and coinage was stated as only a week (Craig 1953, p. xvii), but "long delays" are cited elsewhere for the period prior to the Act of 1666 (Feavearyear 1963, p. 3). In the eighteenth century three months was the norm. In 1817 the delay was only five to six weeks, and reduced further until 1820 by immediate payment for three-quarters of the bullion in bills discountable at fourteen days. In 1829–1831 an even more favorable policy was followed, with depositors receiving three-fourths payment immediately and the remainder after assaying. A two-week delay was stated in 1868, and two to three weeks became the ultimate norm (Seyd 1868, p. 158; Whitaker 1919, p. 503; Cross 1923, p. 376).

The Bank of England Act and Charter, each of 1694, gave the Bank the right to buy or sell bullion.[17] Early on, the Bank offered to buy gold from all comers at a fixed price. From 1717 to 1829 this price was £3 17s. 6d. per ounce of standard gold, and from 1829 onward it was £3 17s. 9d. The latter price was institutionalized in the Bank Charter Act of July 19, 1844, which required the Bank to purchase gold bullion from anyone with its notes at that rate. The Bank was entitled to require the bullion to be assayed and melted into bars at the expense of the seller. The statutory price was a minimum; the Bank could, and sometimes did, offer a higher price.[18]

Therefore gold-bullion owners wishing to dispose of their asset had a choice: the mint or the Bank. The Bank's statutory price was 0.1605% lower than the mint price (£3 17s. 9d. versus £3 17s. 10.5d.); but waiting time (and therefore forgone interest) at the Bank was much less than that at the mint, because only weighing and assay (and possible melting into bars), but not coinage, was required. By the mid-eighteenth century, it became clear that the net return to bullion was greater at the Bank, and no private party went to the mint.[19] In 1817–1820 and 1829–1831 favorable arrangements resulted in use of the mint by some large bankers and bullion brokers. Between these periods private coinage at the mint was minimal, and after 1831 only two such transactions are known. Each transaction was small in amount and, as expected, involved an interest loss beyond the 0.1605-percent price advantage.

18.6 Bank Provision of Bars and Foreign-Exchange Dealings

At some point, probably in the mid-nineteenth century, the Bank began selling gold bars, permitted but not required by its Charter and the Bank of England Acts. The Bank frequently varied its selling price, rather than keeping it steady at the mint price. Winston Churchill's famous Budget Speech of April 28, 1925, announced the introduction of a bill, the Gold Standard Act of 1925, which was passed on May 13 and obliged the Bank of England to sell gold bars to any person at the fixed price of £3 17s. 10.5d. per ounce of standard fineness, the mint price. Payment could be made in any legal tender. A minimum transaction size was specified, 400 ounces of fine gold, which was slightly over £1699 in value.[20]

Unlike the US Treasury, the Bank of England dealt directly in American eagles and other foreign coin, pricing them by weight, an alternative to treating them as bullion. The Bank began this practice in 1852 or possibly earlier (Sayers 1986, pp. 48–49). For many years the Bank set its buying and selling prices of eagles (and other foreign coin) at levels to leave transactors indifferent between eagles and other forms of gold (bars or sovereigns). Beginning in 1890 it made prices more favorable for private parties and also more variable. With the British mint price at 77s. 10.5d. per ounce 11/12ths fine, the equivalent mint price for American gold coin, only 9/10ths fine, was 76s. 5.50909…d. (5 28/55d.) per ounce. Almost always the Bank's purchase price was below, and always its selling price was above, the mint-price equivalent.

18.7 Convertibility of Banknotes and Bank Deposits

Unlike the American experience, banking was in existence in England from the mid-seventeenth century, beginning, and most developed, in London.[21] The most important bank, of course, was the Bank of England, founded in 1694, and which performed the functions of both a commercial and central bank. It had a legal monopoly of joint-stock banking (meaning organization with more than six partners) throughout England (but not Scotland, which had a separate banking system) and an effective monopoly of note issue in and around London.

The earliest banks were the private banks in London, said to be 24 in 1725, 52 in 1785, and nearly 70 in 1800. There were also country banks (located outside London), their number estimated at 12 in 1750, 119 in 1784, and 230 in 1797. Legislation of 1826 permitted joint-stock banking outside 65 miles of London and the Act of 1833 allowed joint-stock banking in London absent note issuance. Bank of England notes served as well as coin as bank reserves, even though the notes did not receive legal-tender status until 1833.

18.8 Legal Versus Effective Metallic Standards

Britain was legally on a silver standard until gold was coined in the mid-thirteenth century, on a bimetallic standard until 1816 (or perhaps 1774, when the legal tender of silver was first limited; or 1798, when free coinage of silver was terminated), and then on a gold standard. Just as for the United States, legal bimetallism translated into effective monometallism. Until the end of the seventeenth century, the British mint ratio was generally below European gold/silver price ratios; so a silver standard was in effect. The annual Soetbeer data exist only from 1687; from that year to 1696 the UK mint ratio (14.49), established in 1663, was uniformly below the Hamburg market ratio.

For a few years at the turn of the eighteenth century, England actually had effective bimetallism: "for a while there were in circulation plenty of full-weight coins of both kinds" (Feavearyear 1963, p. 152). However, foreign gold/silver price ratios had been falling and, after having been increased greatly in 1696, the British mint ratio was not subsequently reduced enough to compensate. At 15.57 the British ratio from 1699 to 1717 was higher than the Hamburg market ratio in all years. It was also higher than the ratio in most European countries—16 in Spain and Portugal, but 15 in France and below 15 in seven other countries—and much higher than the ratio in the Far East, 9–10 in China and Japan and 12 in India.[22]

With gold so overvalued and silver undervalued in England, bimetallism could be only transitory. Arbitrageurs did not take long to adjust to the new situation, exporting silver and importing gold, and Britain shifted to a de facto gold standard. In 1702 silver coinage dwindled to almost nothing, and silver coinage was insignificant in amount in almost every year thereafter until 1816.[23]

Why Britain formally adopted the gold standard in 1816 is controversial. Four explanations have been offered. The conventional view is that Liverpool's Act merely ratified the prevailing effective gold standard that had been in existence since the beginning of the seventeenth century. As Li (1963, p. 174) comments: "It is certain that England did not establish the gold standard by any conscious or deliberate act. Nor was it foreseen by anyone that the gold standard would be established. It was established in practice first and then recognized officially later."

In contrast, Redish (1990) declares that the abandonment of legal bimetallism occurred because of new mint technology and policy, which permitted a gold standard with subsidiary silver coinage. New technology produced coins not readily counterfeitable, and new policy guaranteed the convertibility of all coins at face value. Friedman (1990), however, argues that the achievement of a silver token coinage was not a sufficient reason for the adoption of a gold rather than a silver standard or even for the maintenance of bimetallism when Britain left the then paper standard in 1821. It might have been mere chance that a monometallic gold rather than silver standard replaced the inconvertible pound. David Ricardo, who had great influence, happened to favor a gold standard in 1819. If he had argued rather for silver, that might have been the standard adopted. Britain, with its dominant economic and financial power in the world, might even have made bimetallism work.

Finally, Feavearyear (1963, p. 214) asks what would have transpired under the 1816 Act had there been issuance of the proclamation to permit free coinage of silver at the fixed rate of 62 shillings per pound Troy (with actual coining at 66 shillings). Because of the low market price of silver, massive amounts of silver would have been presented to the mint and a silver standard, albeit with a seigniorage of 6.06%, would have resulted. According to Li (1963, p. 166), the provision for free coinage was based on the assumption that the eighteenth-century experience of a silver market price between 62 and 66 shillings per pound Troy would continue. Instead, the market price fell below 62 shillings. The mint practice of buying all its silver in the open market in lieu of free coinage not only yielded higher profits but also preserved the gold standard.

18.9 Paper Standards

In contrast to the American experience, prior to World War I, Britain had only one episode of a paper standard. On February 27, 1797, acting on government orders issued the previous day (a Sunday), the Bank of England suspended specie payments, that is, it refused to pay out gold for its notes. The note-issuing commercial banks of England and Scotland followed, and the entire country was on a paper standard. The Bank Restriction Act of May 3, 1797 confirmed the suspension, and successive continuing Acts kept it in force until May 1, 1821, when payment resumed, ending the Bank Restriction Period.[24]

With the onset of World War I, the requirement of the Bank of England to redeem in gold coin both its own notes and Treasury currency notes was effectively abrogated by moral suasion, legalistic action, and regulation; and a paper standard resulted.[25] Contrary to the American experience and, ironically, convertibility of paper currency was legally terminated when the country "returned to the gold standard." The Gold Standard Act of 1925 ended both the Bank of England's obligation to pay its notes in coin and the right of holders of currency notes to redemption in coin.[26] The Gold Standard Act also put a legal end to free coinage, permitting use of the mint only by the Bank of England. This provision simply codified a practice that had existed for many years.

Brown (1929, pp. 229–230) writes, with justice: “Thus free interchangeability between Bank of England notes and gold and between gold bullion and gold coin was for the first time legally suspended. And this by the very act that restored England to a gold standard.” Therefore the *international* dollar-sterling gold standard that began in 1925 was complemented by a *domestic* gold standard (involving circulation of gold coin) in the United States but not in Britain. The Bank’s legal obligation to transact in gold pertained exclusively to bars—a form of gold suitable for international movement but hardly for domestic circulation. Table 18.1 summarizes the British monetary standards from 1791 to the present.

The obligation of the Bank of England to purchase bars with its notes at 77s. 9d. per standard ounce was not repealed until the Currency and Bank Notes Act of February 28, 1939 (Public General Acts 1940, pp. 27–28). Its obligation to sell gold bars at the mint price was suspended by government instruction announced on September 20, 1931 (a Sunday) and terminated by the Gold Standard (Amendment) Act the following day.[27] The gold-bullion standard instituted by Churchill’s famous Budget Speech was over. The UK gold standard, running from April 28, 1925 to September 19, 1931, delineated the dollar-sterling international gold standard of the time, because the UK, inner, period was enveloped by

Table 18.1 Effective British monetary standards, 1791 to present

Time period	*Standard*	
	Domestic	*International*
January 1, 1791–February 26, 1797	gold	gold
February 27, 1797–April 30, 1821	paper	paper
May 1, 1821–August 1914	gold	gold
August 1914–April 27, 1925	paper	paper
April 28, 1925–September 19, 1931	paper	gold
September 20, 1931 to present	paper	paper

Sources Brown (1940, pp. 34–35), Cannan (1925, pp. xi, xxxiv), and Yeager (1976, pp. 321–322, 342) the American, which ran from March 18, 1922 to March 5, 1933 (see Chapter 17, Table 17.2).

Divergences from an international gold standard occasioned by World War I were even more extensive for Britain than for the United States, as shown in Chapter 17, Table 17.3. Further, unlike the United States, Britain had a long history of such deviations. Import of bullion was first prohibited by the Act of 1774, but only domestic silver coin and of less than standard fineness was affected. The purpose was to restrict supply, thus maintaining the value of silver coins in spite of light weight and the refusal of the government to re-coin them. In fact, the silver coinage was so clipped, sweated, and naturally worn that it became no longer profitable to melt it down and sell it at the market price for silver bullion. An, albeit unsatisfactory, subsidiary silver coinage was in existence even before the currency reform instituted by the Coinage Act of 1816.

From early in World War I to April 1919, gold imports were restricted extra-legally by the simple expedient of the Bank of England purchasing all imported gold. This was accomplished principally by arrangement with South Africa to sell all gold to the Bank at the price of 77s. 9d. per standard ounce. From December 5, 1916 to March 31, 1919 imports were prohibited by Royal Proclamation except for gold sold to the Bank. Beginning September 12, 1919, gold could again be freely imported and sold in London at the market-determined price.

Restriction of export of gold and silver goes back to the turn of the thirteenth century; but the restraints were often evaded. "Act after Act and proclamation after proclamation attempted to prevent the export of coin and, later, of precious metal in any form…For a long period death was the penalty for those found exporting good English money" (Feavearyear 1963, pp. 3–4). The king could override any export prohibition by license. Such a license was required to export gold or silver of any kind until the Act of August 1, 1663, which permitted the free export of gold and silver bullion and the re-export of foreign coin.

An Act of 1696 required export of bullion to be stamped and an oath taken that it was not produced from domestic coin. Of course, lying permitted melted British coin to be exported under the Act. Peel's Act of 1819 repealed all restrictions on the melting and exporting of British gold coin and all but one for silver. The remaining restriction, concerning taking an oath that bullion to be exported was not produced from clippings of silver coin, was removed by an Act of 1821.

During World War I, gold exports were restricted solely by extralegal means (apart from Trading with the Enemy legislation), not only by the Bank of England commandeering all gold imports but also by effective moral suasion of bankers and bullion brokers. The London bullion market suspended operations (Brown 1929, p. 6; 1940, pp. 36, 1014; Cross 1923, pp. 377–378; Jaeger 1922, p. 21). On April 1, 1919, just when imports were liberalized, an Order in Council under the Customs (Exportation Prohibition) Act of 1914 was issued forbidding the export of gold coin and bullion except by license. Beginning September 12, 1919, the Treasury agreed to provide licenses for the re-export of gold, and the London gold market reopened.

The Gold and Silver (Export Control, etc.) Act of December 23, 1920 prohibited the export of gold with the important exception of gold produced within the British Empire and imported under any arrangement approved by the Treasury. The Act was to continue in force until December 31, 1925. Churchill's Budget Speech of April 28, 1925 superseded the Act by giving the Bank of England a general license to export gold, which was in effect an obligation to provide bars to private exporters.

At the beginning of August 1914, by proclamations, international (and domestic) payments were postponed—the so-called "Moratorium." This embargo on payments, which lasted until November 4, constituted an extreme type of exchange control and, because of the importance of London in financing international trade and payments, made the international gold standard non-operational.[28] Conventional exchange control was in existence in England from the twelfth to the seventeenth century, varying in intensity and only partially effective (Einzig 1970, pp. 104–108, 157–160).

With World War I, exchange control took several forms. First, beginning July 1915, there was a mobilization of foreign securities, at first by taxation and moral suasion, ultimately by order. Second, in 1916 and 1917, regulations were promulgated that prohibited the export of capital in any form; they were removed in August 1919. Third, beginning December 24, 1914, overseas investment taking the form of new issues of capital was tightly controlled by the Treasury. From November 1919 the control of new issues was done through moral suasion on the part of the Bank of England. Removed, reinstituted, and removed again in 1924–1925, the control was reintroduced in 1929, relaxed and then re-imposed in 1930.

As a result of strong moral suasion, the Bank of England achieved even greater impounding of domestic gold holdings in World War I than did the US Treasury. Both commercial banks and the public exchanged their coin for Bank of England notes.[29]

As mentioned above, until 1819 British coin could not legally be melted into bullion for the purpose of export. During World War I, on December 5, 1916, a regulation under the Defense of the Realm Act prohibited the melting of gold coin or its use other than as currency. This restriction continued under the Gold and Silver (Export Control, etc.) Act of 1920, which provided for exceptions upon license from the Treasury. On May 18, 1918 a regulation prohibited buying or selling gold coin at a premium in Britain, which restriction may in effect have been incorporated in the 1920 Act under the phrase "use otherwise than as currency." The purpose of both regulations, absent from the American experience, was to remove any incentive to redeem Bank and currency notes.

Britain also engaged in intervention in the dollar-sterling exchange market in World War I. From August 1915 to January 12, 1916 intervention took the form of exchange-rate management. This practice was followed by exchange-rate pegging, at first between \$4.765 and \$4.77, after May 1916 between \$4.764375 (\$4.76 7/16) and \$4.765625 (\$4.76 9/16) per pound. Intervention continued after the war, ending in March 1919. So, during 1914–1925, when an international gold standard did not exist, the dollar-sterling exchange rate did not float freely until almost half the period was over.

Notes

1. Discussion of the British unit of account is based on the excellent presentation in Feavearyear (1963, pp. 6–9).
2. The term "pound sterling" emanates from the fact that in the twelfth century the penny was called a "sterling," possibly emanating from *steorra* (Latin for "star"), some of the early coins having been adorned with a star. For alternative speculation, see Craig (1953, p. 6).
3. Useful histories of the British monetary standard are provided by Feavearyear (1963), Craig (1953), and, though with limited time spans, Ashton (1955, pp. 167–177), Horsefield (1960), Horton (1887), and Li (1963).
4. These Acts are in *International Monetary Conference* (1879, pp. 345–347, 373–378).

5. The order and Royal Proclamation for coinage of the guinea and sovereign, respectively, are in Horton (1887, pp. 229–230, 282–283).
6. The Order is in Horton (1887, pp. 229–30).
7. The relevant sections of the Acts are in Horton (1887, pp. 243–246).
8. The Treasury order and the report adopted by Parliament are in Horton (1887, pp. 250–254).
9. The Proclamation is in *International Monetary Conference* (1879, p. 316).
10. For the Proclamation, see Horton (1887, pp. 282–283).
11. The Act is in *Public General Statutes* (1870, pp. 153–162).
12. See Horton's excellent discussion in *International Monetary Conference* (1879, pp. 373–374).
13. These Acts are in *International Monetary Conference* (1879, pp. 332–349).
14. The Resumption Act is fully described in Feavearyear (1963, p. 221). The relevant section of the Bank of England Act, 1833, and the Currency and Bank Notes Acts are in Shrigley (1935, pp. 40, 62–65, 79–85).
15. The 1666 and 1768 Acts are in *International Monetary Conference* (1879, pp. 309–314).
16. See the tables in Feavearyear (1963, pp. 435–436).
17. For the relevant parts of these documents, see Shrigley (1935, pp. 1, 27).
18. For the relevant section of the Bank Charter Act, see Shrigley (1935, pp. 43–44).
19. Of course, the Bank coined its bullion at the Mint as the Mint's only customer. It would even do so at a loss, to preserve the gold-convertibility of its notes.
20. The relevant portion of the Budget Speech and the entire Gold Standard Act are in Sayers (1976, pp. 80–86).
21. For the early history of banking in Britain, see Cameron (1967, pp. 15–99), Clapham (1945, vol. 1, pp. 156–172), and Kindleberger (1993, pp. 53–6, 77–96).
22. These foreign ratios were stated by Isaac Newton, Master of the mint, in a report of September 21, 1717. The report is in *International Monetary Conference*, pp. 317–318, and the ratios are conveniently summarized in Li (1963, pp. 151–152).
23. Annual data on coinage of gold and silver are in Craig (1953, pp. 410–421).
24. Good discussions of the Bank Restriction Period are in Cannan (1925, pp. vii–xlvi) and Kindleberger (1993, pp. 63–66).
25. See Cross (1923, p. 377), Brown (1929, p. 6; 1940, pp. 34–35), and Fraser (1933, pp. 32–33).
26. The decision to end legal redemption was due to a breakdown of the moral suasion of private bankers to deny gold to their customers. See Sayers (1986, pp. 147–148).

27. The press notice announcing the order is in Sayers (1976, pp. 264–265) and the Act is in Shrigley (1935, p. 86).
28. For a history of the Moratorium and texts of the documents, see Kirkaldy (1921, pp. 1–14, 405–409).
29. See Beckhart (1924, p. 252), Shrigley (1935, p. 63, n. 1), and Brown (1940, p. 42).

References

Ashton, Thomas S. *An Economic History of England: The 18th Century*. London: Methuen & Co., 1955.

Beckhart, Benjamin Haggott. *The Discount Policy of the Federal Reserve System*. New York: Henry Holt and Company, 1924.

______. *The New York Money Market, vol. Ill, Uses of Funds*. New York: Columbia University Press, 1932.

Brown, William Adams, Jr. *England and the New Gold Standard 1919–1926*. London: P. S. King & Son, 1929.

______. *The International Gold Standard Reinterpreted*, 2 vols. New York: National Bureau of Economic Research, 1940.

Cameron, Rondo. *Banking in the Early Stages of Industrialization*. New York: Oxford University Press, 1967.

Cannan, Edwin. *The Paper Pound of 1797–1821*. London: P. S. King & Son, 1925.

Clapham, Sir John. *The Bank of England*, 2 vols. Cambridge: Cambridge University Press, 1945.

Craig, Sir John. *The Mint: A History of the London Mint from A.D. 287 to 1948*. Cambridge: Cambridge University Press, 1953.

Cross, Ira B. *Domestic and Foreign Exchange*. New York: The Macmillan Company, 1923.

Einzig, Paul. *The History of Foreign Exchange*, 2nd edn. London: Macmillan and Co., 1970.

Feavearyear, Sir Albert. *The Pound Sterling*, 2nd edn. Oxford: Clarendon Press, 1963.

Fraser, Herbert F. *Great Britain and the Gold Standard*. London: Macmillan and Co., 1933.

Friedman, Milton. "Bimetallism Revisited." *Journal of Economic Perspectives* 4 (Fall 1990), pp. 85–104.

Horton, S. Dana. *The Silver Pound*. London: Macmillan and Co., 1887.

Horsefield, J. Keith. *British Monetary Experiments. 1650–1710*. Cambridge, MA: Harvard University Press, 1960.

International Monetary Conference. Senate Executive Document No. 58, 45th Cong., 3rd session. Washington, DC: Government Printing Office, 1879.

Jaeger, Ruth M. *Stabilization of the Foreign Exchanges*. Ph.D. dissertation. New York: Columbia University, 1922.

Kindleberger, Charles P. *A Financial History of Western Europe*, 2nd edn. New York: Oxford University Press, 1993.

Kirkaldy, Adam W. *British Finance During and After the War, 1914–21.* London: Str Isaac Pitman & Sons, 1921.

Li, Ming-Hsun. *The Great Recoinage of 1696 to 1699.* London: Weidenfeld and Nicolson, 1963.

Public General Statutes, 1870, vol. V. London; Queen's Printers, 1870.

Public General Acts, 1938–39, vol. I. London: King's Printer, 1940.

Redish, Angela. "The Evolution of the Gold Standard in England." *Journal of Economic History* 50 (December 1990), pp. 789–805.

Sayers, Richard S. *The Bank of England 1891–1914.* Appendixes. Cambridge: Cambridge University, Press, 1976.

______. *The Bank of England 1891–1914.* Cambridge: Cambridge University Press, 1986.

Seyd, Ernest. *Bullion and Foreign Exchanges*. London: Effingham Wilson, 1868.

Shrigley, Irene. *The Price of Gold*, London: P. S. King & Son, 1935.

Spalding, William F. *Dictionary of the World's Currencies and Foreign Exchanges*. London: Sir Isaac Pitman & Sons, 1928.

Whitaker, Albert C. *Foreign Exchange*. New York: D. Appleton and Company, 1919.

______. *Foreign Exchange*, 2nd edn. New York: D. Appleton-Century Company, 1933.

Yeager, Leland B. *International Monetary Relations: Theory, History, and Policy*. New York: Harper & Row, 1976.

CHAPTER 19

Afterword to Part V

Chapters 17 and 18 provide detailed histories of the U.S. and British monetary standards, thus extending the over-all discussion of monetary standards in Part IV.[1] In turn, Chapters 17–18 serve as background to the analysis of dollar-sterling exchange-market efficiency (Chapter 21).

As Alan M. Taylor (1998) observes:

> Part One of the book [Officer, 1996] lays down the key historical and institutional features of the landscape from the beginning of the dollar-sterling gold standard in 1791 (when the U.S. went to a formal metallic standard) to its demise in 1931 (when Britain suspended convertibility). The laws and mechanics of coinage, minting, convertibility of paper to metal, dealings in the market and at banks, and so forth are all carefully described. The text and tables note significant legislative acts forcing regime changes for both countries in this entire time span, including changes in the metal of the standard for the U.S., and changes in parities for both countries (i.e., the metal content of the unit of account). Periods of convertibility and inconvertibility are shown.

Similarly, Marc Flandreau (1998, p. 1223) states: "the author covers in minute detail the evolution of monetary standards in the United States and United Kingdom."

The history of U.S. and British monetary standards is summarized in Table 19.1. Pertinent to Part VI are the implications for the dollar-

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_19

Table 19.1 U.S. and British monetary standards

Period	*Domestic standard*[a,b]	
	United States	*Britain*
January 1, 1791–February 26, 1797	silver	gold
February 27, 1797–August 29, 1814	"	paper
August 30, 1814–February 19, 1817	paper	"
February 20, 1817–April 30, 1821	silver	"
May 1, 1821–July 30, 1834	"	gold
July 31, 1834–May 9, 1837	gold	"
May 10, 1837–May 9, 1838	paper	"
May 10, 1838–October 9, 1839	gold	"
October 10, 1839–March 17, 1842	paper	"
March 18, 1842–October 13, 1857	gold	"
October 14–December 13, 1857	paper	"
December 18, 1857–December 29, 1861	gold	"
December 30, 1861–December 31, 1878	paper	"
January 1, 1879–August 1914	gold	"
August 1914–April 5, 1917	"	paper
April 6, 1917–March 17, 1922	paper	"
March 18, 1922–April 27, 1925	gold	"
April 28, 1925–September 19, 1931	"	paper[c]
September 20, 1931–March 5, 1933	"	paper
March 6, 1933–January 30, 1934	paper	"
January 31, 1934–August 14, 1971	paper[c]	"
August 15–December 17, 1971	paper	"
December 18, 1871–February 12, 1973	paper[c]	"
February 13, 1973–	paper	"

[a]Where international standard differs, specified in footnote
[b]Exceptional standards omitted. See source for details
[c]Gold, international standard
Source Table 17.2, chapter 17, and Table 18.1, chapter 18

sterling exchange rate. When both countries are on a metallic standard, the dollar/sterling exchange rate is "fixed," meaning that it has an upper bound and lower bound (the "gold points" or "specie points"). Much of the 1791–1914 period is of that ilk.

When both countries are on a paper standard and neither country intervenes in the foreign-exchange market; or when one country is on a metallic and the other on a paper standard, with the latter country not intervening: then the exchange rate is freely floating. That is a reasonable characterization of the dollar-sterling market during 1797–1821, 1837–1842, 1857, and 1862–1878—encompassing all paper standards prior to World War I.

The later paper-standard periods encompass episodes of free floating (1920–1924, 1931–1932), managed floating (1932–1939, 1971, 1973–), intervention points (1925–1931, 1947–1971, 1971–1973), pegged rate via exchange control and other restrictions (1914–1919, 1939–1947).[2]

Notes

1. In addition, Officer (undated) offers a quantitative history of British and American gold prices (both market and official) from their origins to the present, with the data readily retrievable in Officer and Williamson (2021).
2. Month designations are omitted from the episodes. More importantly, some periods involve more than type of episode (for example, exchange control persisting beyond the stated end-of-periods). I classify episodes according to the main genre.

References

Flandreau, Marc (1998). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. *Economic Journal* 108 (July), pp. 1223–1225.

Officer, Lawrence H. (1996). *Between the Dollar-Sterling Gold Points*. Cambridge: Cambridge University Press.

——— (undated). *What Was the Price of Gold Then? A Data Study*. Measuring-Worth website.

———, and Samuel H. Williamson (2021). *The Price of Gold, 1257–Present*. MeasuringWorth website.

Taylor, Alan M. (1998). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. EH.net.

PART VI

Monetary-Standard Behavior

CHAPTER 20

Bullionist Periods

20.1 Bullionist Controversies (Empirical Evidence)

Originally published in *The New Palgrave Dictionary of Economics*, Third edition. Palgrave Macmillan, 2018, pp. 1150–1156.

The bullionist periods of Sweden, England, and Ireland involved bullionist versus anti-bullionist macroeconomic debates, with empirical studies vindicating largely the anti-bullionist side.

20.1.1 History of Bullionist Periods

The bullionist controversy is a debate that can occur in monetary history when a paper currency and floating exchange rate interrupt a metallic standard. The three famous bullionist periods pertain to Sweden, England and Ireland. In 1745, the Riksbank made its notes inconvertible into copper bullion, resulting in the paper daler. It was not until 1776 that the Swedish bullionist period ended, with conversion to a new currency unit (the riksdaler) on a silver standard. The English, followed by the Irish, bullionist period began in 1797, each by government order requiring the Bank of England and Bank of Ireland to cease making gold payments for its notes. Legislation, periodically renewed, solidified the orders. In 1821

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_20

the Bank of England, followed by the Bank of Ireland, resumed payment in gold, and the countries were back on a gold standard. The English episode is called the 'Bank Restriction Period'.

The three bullionist periods involved common elements: a prior metallic standard replaced by a paper standard, a fixed exchange rate (constrained within a band around an effective mint parity) giving way to a floating rate, unusually high inflation, depreciation of the currency in the foreign-exchange and bullion markets, a subperiod of deflation, and eventual return to a specie standard and fixed exchange rate. Also, periods of war occurred both before and during the bullionist periods.

Some characteristics were shared by only two of the periods. First, the proximate cause of the Swedish and English Restrictions was a tremendous loss of reserves on the part of the Riksbank and Bank of England. This was not the case for the Bank of Ireland; British pressure induced the Irish government to suspend convertibility of Bank of Ireland notes. Second, for Sweden and England, their main trading partners remained on a metallic standard. This was not so for Ireland, with England also on paper. Third, England and Ireland returned to a gold standard at the old parity; Sweden switched from an effective copper to an effective silver standard, and banknotes were depreciated by 50% in terms of silver.

Two additional features characterize all three periods. First, the macroeconomic debate centered on determination of the exchange rate and price level, and their relationship to the balance of payments and note issues of the central bank. The bullionists adopted a monetarist approach, and the anti-bullionists a non-monetarist position. Second, Parliament played a key role in the controversy. In the case of Sweden, two political parties vied for control of Parliament. The 'Caps' had a bullionist agenda, and the 'Hats' an anti-bullionist policy. Both had intellectual supporters on the outside. The British House of Commons appointed committees, in 1804 and 1810, to investigate the depreciated Irish and English currencies. Each committee produced a highly bullionist report, important in the literature; but in neither case was the report favorably received by Parliament.

20.1.2 *Bullionist, Anti-Bullionist, and Country-Bank Models*

To examine the empirical literature on the bullionist controversies, each side is represented by its mainstream model of chains of causality, sequential hypotheses. Notation is X $\rightarrow$ Y ('X causes Y, with $\partial Y/\partial X > 0$').

Multiple hypotheses are W, X $\rightarrow$ Y ('W $\rightarrow$ Y and X $\rightarrow$ Y') and X $\rightarrow$ Y, Z('X $\rightarrow$ Y and X $\rightarrow$ Z'). The subscript f designates a foreign variable. Variables are:

BN: central-bank notes in circulation
BP: balance-of-payments deficit
CN: country banknotes in circulation
ER: exchange rate, price of foreign currency
FR: remittances to foreign countries
HQ: quantity and quality of harvest
MS: money supply (M1)
PG: price of gold
PL: price level
PM: price of imports
PW: price of wheat
TR: foreign trade-restrictions

The *bullionist model* is decidedly monetarist: only monetary variables affect only monetary variables. The English-bullionist chain of causation is:

$$\text{BN} \rightarrow \text{MS} \rightarrow \text{PL} \rightarrow \text{ER, PG}$$

BN $\rightarrow$ MS reflects the bullionist, and correct, perception that Bank of England notes constituted the monetary base during the Restriction Period. There was a hierarchy of banks: Bank of England (central bank), London private banks, and country banks. Bank of England notes (held as reserves by the country banks and London private banks) were non-redeemable; deposits at the Bank (held as reserves only by the London private banks) were cashable only in Bank of England notes. The country banks—but not the London private banks—issued notes. There were no legal reserve requirements for any bank; but, like all companies, banks had to settle their debts (note and deposit liabilities) in cash. Reserves of the country banks were principally deposits at the London private banks, with Bank of England notes (and, in principle, gold) for vault cash. Bank of England notes circulated in and around London, as well as in Lancashire and Norwich; country banknotes circulated elsewhere in England and Wales. During the Bank Restriction Period, the English country banks and Scottish banks 'redeemed' their notes in Bank of England notes rather than gold. This was a matter of practice rather than law.

Strictly speaking, gold coin was a component of the monetary base, but the premium on gold bullion did not have a counterpart in the premium of gold coin over Bank of England notes. There was no legal market for domestic coin in terms of paper money, and an overwhelming proportion of the gold coin nominally in circulation or newly minted was in fact hoarded or exported.

For the bullionists (and anti-bullionists), the money supply had as components Bank of England notes, country banknotes, and coin. In excluding *deposits* from M1, the writers of the Restriction Period were not far off the mark. First, except in London, 'deposits' generally meant time or savings deposits rather than demand deposits. Second, if interbank transactions are excluded, demand deposits typically were exchanged for cash rather than transferred to another account.

BN $\rightarrow$ MS was also asserted by the Irish bullionists, even though the banking system was looser. In and around Dublin, notes of the Dublin private banks circulated along with notes of the Bank of Ireland. Gold did not circulate, except in the north until 1808–1809, when it was replaced by the notes of newly established Belfast banks. Elsewhere, local private banknotes generally dominated, but in competition with Bank of Ireland notes and, to a lesser extent, Dublin private-bankers' notes. The private banks kept their reserves in Bank of Ireland notes (and gold), and by convention their notes were redeemed in Bank of Ireland notes.

In the Swedish bullionist period, BN = MS. With little coin circulating, no commercial banks in existence, and deposits at the Riksbank representing merely the right to make withdrawals in notes, Riksbank notes essentially equaled the money supply.

MS $\rightarrow$ PL pertains to the quantity theory of money. Underlying this theory is the bullionist view that the Bank of England effectively pegged the market interest rate at five percent, by standing ready to discount all 'good' commercial bills at that rate. Thus the monetary base is perfectly elastic at the constant discount rate of five percent, a powerful impetus to the quantity theory.

There is good reason for this view: the usury laws set a five-percent limit on annual interest on bills of exchange, and the discount rate of the Bank of England was fixed at this rate. While bill brokers could charge a commission and private banks could require a minimum balance, the Bank did not use such devices. The market discount rate (for good bills) did not *exceed* five percent during the Restriction. In fact, only for about a year (beginning July 1817) did the market rate even *fall below* five percent. The situation was yet stronger regarding the Bank of Ireland: its discount rate was limited to five percent by charter.

However, the English and Irish bullionists were wrong in inferring that the monetary base (essentially BN) could rise without limit. First, there is evidence that in historical fact the monetary base was not perfectly elastic. Only 'good' bills—a minority of bills—were acceptable by the Banks. Also, the Bank of England effectively regulated discounts via a rationing system. These facts act against the quantity theory but support the concept of BN as an autonomous policy variable.

Second, even if the *supply* of the monetary base (essentially BN) is perfectly elastic at the pegged market interest rate, BN is limited by the *demand* for the monetary base. The Bank of England and Bank of Ireland could not induce the private sector to hold more BN than demanded. BN was viewed by the bullionists as the first link in the causal chain; but it is an endogenous variable. A low level of economic activity could hold down the demand for BN.

PL $\rightarrow$ ER is the purchasing-power-parity theory (given PL_f), the causal nature of which is generally ignored in the modern literature. PL $\rightarrow$ PG involves a relatively unchanged PG_f, for, under perfect markets, PG is the product of ER and PG_f. PG was not as interesting to the Swedish and Irish bullionists as it was to the English. Sweden had been on a copper standard; the concern in Ireland was depreciation of the Irish currency against the British. For the Swedish and English protagonists, foreign exchange was Continental currencies.

For most Swedish and Irish bullionists, the latter part of the chain is merely MS $\rightarrow$ PL, ER. The price level and exchange rate are co-determined by the money stock. Some Irish bullionists allowed for a changing foreign (English) price level, so the hypothesis becomes MS/MS_f (or BN/BN_f) $\rightarrow$ ER.

The English *anti-bullionist model* involves a balance-of-payments theory of the exchange rate, with demand for and supply of bills of exchange represented by the payments deficit (BP), yielding ER and PG. The state of the harvest, a real factor, determines the domestic price of grain, represented by the price of wheat (PW). The exchange rate is an ingredient in the price of imports, which, together with PW, determines PL. These anti-bullionists saw three principal determinants of BP, that is, of shifts in the demand for or supply of foreign exchange: PW, foreign trade-restrictions (wartime restraints: the Continental System and the American embargo), and foreign remittances (external government payments: direct military expenditure and subsidies to allied countries). The English anti-bullionist causal chain is:

$$\begin{tikzcd}
1/HQ \arrow[r] & PW \arrow[rr] \arrow[d] & & PL \arrow[r] & BN \\
TR, FR \arrow[r] & BP \arrow[r] & ER, PG \arrow[r] & PM \arrow[u] &
\end{tikzcd}$$

In emphasizing the price of wheat, the anti-bullionists recognized the highly agrarian state of the British economy, notwithstanding the industrial revolution in progress. The emphasis on wartime interference with trade and on external military expenditure reflected the French Revolutionary and Napoleonic Wars, in which Britain was engaged for much of the Bank Restriction Period.

For the Irish anti-bullionists, concerned with the English exchange, TR and PG were unimportant. They did not make explicit the connection of PW and PM to PL, and FR took the form of payments to absentee landlords in England. Some consolidated the trade balance, interest payments, net capital exports, and FR, to compose (and presumably shift) BP in the causal chain. They left unclear the mechanism from BP to PL. The Swedish anti-bullionists had the chain:

$$BP \rightarrow ER \rightarrow PM \rightarrow PL, \text{ allowing real shocks to operate on BP.}$$

The anti-bullionists used the 'real-bills' doctrine to reverse the bullionist BN $\rightarrow$ PL causation. They accepted that the Bank behaved passively in its note issuance, but used the real-bills theory to demonstrate that excess issue (beyond the 'needs of trade') would be returned to the Bank instead of acting to increase the price level monetarily. Only non-monetary forces could cause real income and then the price level to increase, and would underlie the demand for discounting to finance a higher volume of transactions, whence PL $\rightarrow$ BN. The Irish bullionists also propounded the real-bills doctrine (for the Bank of Ireland), although some saw ER playing the role of PL.

Bullionists in all three periods essentially inverted the real-bills theory by offering the *policy rule* that central-bank note issuance should be oriented to the exchange rate and (for the English bullionists) gold price: ER, PG $\rightarrow$ 1/BN.

20.1.3 Extension to Country Banks

A subsidiary part of the English and Irish bullionist controversies was the extent to which the country banks (in Ireland, including Dublin private banks) could affect the money supply independent of the central bank. Should the first hypothesis in the bullionist chain, BN $\rightarrow$ MS, incorporate CN naturally as BN $\rightarrow$ CN $\rightarrow$ MS (country banks unable to vary their note issues independent of the central bank)? Or should the hypothesis be (BN + CN) $\rightarrow$ MS (the central bank and country banks able either jointly or separately to change their issues)? Or should the hypothesis be CN $\rightarrow$ MS (only the country banks, not the central bank, having the power to change the money supply)? The question was answered differently by groups that cut across the bullionist versus anti-bullionist line.

The correct hypothesis is not clear, because of the environment in which banks operated. Among the complicating, and largely unknown, elements are the extents to which (a) one-time replacement of gold by central-bank notes in reserves altered country-bank policy regarding reserve ratios, (b) country-bank reserve ratios varied over time, (c) public preference for central-bank over country-bank notes changed in particular geographic areas and over time, (d) circulation of counterfeit notes and unlicensed-bank notes affected the demand for and supply of country-bank and central-bank notes, and (e) London private banks were prepared to run down their reserve ratios to accommodate country-bank demand for additional reserves.

20.1.4 Empirical Studies: Visual Comparison of Movements of Variables

The empirical studies examined here make use of quantitative information to test one or more component hypotheses of the bullionist or anti-bullionist models. It is logical to begin with contemporary studies, as it is the hypotheses of contemporary authors that are delineated in the previous sections.

All contemporary investigations use a simple technique: visual inspection of sets of figures, formal tables, or charts. The earliest such studies pertain to the Irish bullionist period, with BN and BN_f the note circulations of the Bank of Ireland and Bank of England. Parnell (1804), Foster (1804), and the 1804 Currency Report (in Fetter 1955) find that BN $\rightarrow$

ER is confirmed. Ó Gráda (1993) and Fetter (1955) criticize the Report for its small number of observations and selective observations. These criticisms can be extended to Parnell, but not to Foster. The report of 1804 and Parnell also claim successful testing of $BN/BN_f \rightarrow ER$. Ó Gráda (1991) finds this part of the Report misleading in several respects; but the Report is to be commended for making specific allowance for the replacement of gold coin by notes. The Report also claims to disprove BP → ER, via computation of a net balance-of-payments surplus. However, this proves little, because there is no representation of shifts in the demand for or supply of bills on London.

Contemporary empirical work on the English bullionist period begins with Ricardo (1811), whose positive finding of BN → ER (Hamburg exchange) is reinforced by observation of a lagged effect and by accounting for replacement of gold coin by Bank of England notes. Galton (1813) confirms that BN → ER, PG. Anonymous (1819) sees mixed evidence for that hypothesis, but observes that grain imports and FR (not precisely defined) affect the exchange rate—the first results in favor of anti-bullionism.

There is a hiatus of more than a century, but three groupings of subsequent work do not merit review. First is any investigation, such as Silberling (1924), involving the London price of the Spanish dollar to represent the exchange rate. That choice is methodologically unsound. Britain was on a suspended gold (not silver) standard, and the Spanish silver dollar was not a circulating coin in Hamburg, the main foreign-exchange market. Second are tests making use of Silberling-developed series of Bank of England total advances and their private versus public components. These series have been shown to be seriously inconsistent with the Bank's published data. Third, and most unfortunate, are all studies using 'data' on country banknote circulation. There exist no true data on country banknote circulation in England, or private banknote circulation in Ireland, during the bullionist period. Further, with no legal or fixed reserve ratio of note liabilities to cash, the circulation of the Bank of England, or Bank of Ireland, cannot be used to infer that of the private banks.

Private banks were required to register at the Stamp Office and pay a stamp tax on notes, prior to issuance. Some have used stamp-tax data to develop proxy CN series for England, based on the value of country banknotes stamped; but the series are based on assumptions so tenuous as to make the series unusable.

Silberling (1924) develops an annual series for FR ('extraordinary foreign payments'), consisting of grain imports over a normal amount, Continental British war expenditures, and subsidies to foreign states. Using various definitions of FR, based largely on Silberling, Angell (1926) shows that FR → ER, but can find no causal relationship between PL and ER. This result, favorable to anti-bullionism, is supported by Morgan (1939, 1943) and Viner (1937). Morgan rejects BN → PL, but accepts PL → BN. His only finding not supportive of anti-bullionism is the lack of a relationship between PW and PL or BN.

Gayer et al (1953, p. 932) support BP → ER; but they represent BP by the balance of trade, the data of which are crude. For the Swedish period, Eagly (1971) and Bernholz (1982, 2003) support BN → PL, ER, favorable to bullionism.

This entire body of literature must be viewed with caution. First, interpretation of relationships among variables is subjective when data are merely tabulated or plotted. Second, macroeconomic variables are generally non-stationary, leading to the possible outcome of 'spurious regression'.

20.1.5 Empirical Studies: Time-Series Analysis

Myhrman (1976) computes annual growth rates of BN and PL, for Sweden and England, and argues that BN → PL. Jonung (1976) does the same for Sweden alone. Transforming data to growth rates could yield stationarity. In a joint test of bullionist and anti-bullionist hypotheses, Arnon (1990) regresses PL on PW, BN, and a trend. He finds that BN contributes more to the regression than PW. The variables are transformed to correct for serial correlation, which could correct spurious regression.

Formal time-series analysis in the bullionist literature begins with Ó Gráda (1989, 1993). For England, he cannot reject a cointegration relationship between logPL and logBN. This means that there is no long-term equilibrium between the variables, a failure of support for either bullionism or anti-bullionism The same negative result holds for Ireland, with BN/BN_f used in place of BN.

Nachane and Hatekar (1995) use Granger causality and cointegration techniques for England. Their variables are PL, ER, PG, BP, and BN/Y (transformed to logarithms except for BP, the only non-stationary variable), where Y is real output. Their results are ER → PL, PL →

BN/Y (with PL and BN/Y the only cointegrated pair of variables), and BP → ER, PG. The findings are strongly supportive of anti-bullionism; but measuring the money supply in relation to output is outside the mainstream controversy.

The analyses of Ó Gráda and Nachane-Hatekar are restricted to bivariate econometrics. Officer (2000) applies multivariate testing to PL, ER, BN, FR, and PW, for England. Non-stationarity cannot be rejected, but cointegration is rejected. The logarithmic variables are first-differenced (to achieve stationarity), and Granger causality testing along with innovation analysis is applied. Results are mixed for bullionism, but unambiguously favorable to anti-bullionism. For example, the real-bills doctrine, PL → BN, receives stronger support than does the quantity theory, BN → PL.

It is logical that the time period for testing hypotheses be strictly within the pertinent bullionist period, because the alternative (bullionist versus anti-bullionist) models are geared to a paper standard and floating exchange rate. As his sample, Officer uses the 96 quarters encompassed by the Bank Restriction Period (1797-Q2 to 1821-Q1, where Q is the quarter-year). Nachane and Hatekar employ annual data, and extend the time period to 1838. Ó Gráda has quarterly observations, but begins his time periods prior to 1797.

Nachane and Hatekar can also be criticized for using the exchange rate on Paris rather than Hamburg to represent ER. There are no quotations on Paris until 1802 (whence they lose observations), and historians agree that the Hamburg exchange was more representative during wartime.

To conclude: certainly, at least for England, the anti-bullionist position receives greater support (or less contradiction) than the bullionist side of the controversy. This result is inconsistent with modern macroeconomics. The anti-bullionist approach to the exchange rate (a flow theory) and monetary policy (passive, accommodating the price level) has been superseded in modern theory. Also, modern monetarism emanates from bullionism.

References

Angell, J.W. 1926. *The theory of international prices*. Cambridge, MA: Harvard University Press.

Anonymous. 1819. Two tables...showing the rates of exchange on Hamburgh compared with the amount of bank notes, and the price of gold, and with the foreign expenditure, and the value of grain imported from the year 1793 to 1819. *Pamphleteer* 15: 281–286.

Arnon, A. 1990. What Thomas Tooke (and Ricardo) could have known had they constructed price indices. In *Keynes, macroeconomics and method*, ed. D.E. Moggridge. Aldershot: Edward Elgar.

Bernholz, P. 1982. *Flexible exchange rates in historical perspective*. Princeton: International Finance Section, Princeton University.

______. 2003. *Monetary regimes and inflation*. Cheltenham: Edward Elgar.

Eagly, R.V. 1971. *The Swedish bullionist controversy*. Philadelphia: American Philosophical Society.

Fetter, F.W. 1955. *The Irish pound, 1797–1826*. Evanston: Northwestern University Press.

Foster, J.L. 1804. *An essay on the principles of commercial exchanges*. London: J. Hatchard.

Galton, S.T. 1813. *A chart, exhibiting the relation between the amount of Bank of England notes in circulation, the rate of foreign exchanges, and the price of gold and silver bullion and of wheat*. London: J. Johnson.

Gayer, A.D., W.W. Rostow, and A.J. Schwartz. 1953. *The growth and fluctuation of the British economy, 1790–1850*. Oxford: Clarendon Press.

Jonung, L. 1976. Money and prices in Sweden, 1732–1972. *Scandinavian Journal of Economics* 78: 40–58.

Morgan, E.V. 1939. Some aspects of the bank restriction period, 1797–1821. *Economic History* 4: 205–221.

———. 1943. *The theory and practice of central banking, 1797–1913*. Cambridge: Cambridge University Press.

Myhrman, J. 1976. Experiences of flexible exchange rates in earlier periods: Theories, evidence, and a new view. *Scandinavian Journal of Economics* 76: 169–196.

Nachane, D.M., and N.R. Hatekar. 1995. The bullionist controversy: An empirical reappraisal. *Manchester School of Economic and Social Studies* 63: 412–425.

Officer, L.H. 2000. The bullionist controversy: A time-series analysis. *International Journal of Finance & Economics* 5: 197–209.

Ó Gráda, C. 1989. *The paper pound of 1797–1821: A co-integration analysis*. Working paper. Dublin: Centre for Economic Research, University College.

———. 1991. Reassessing the Irish pound report of 1804. *Bulletin of Economic Research* 43: 5–19.

———. 1993. The Irish paper pound of 1797–1820: Some cliometrics of the bullionist debate. *Oxford Economic Papers* 45: 148–156.

Parnell, H. 1804. *Observations upon the state of currency in Ireland*. Dublin: M. N. Mahon.

Pressnell, L.S. 1956. *Country banking in the industrial revolution*. Oxford: Clarendon Press.
Ricardo, D. 1811. The high price of bullion. In *The works and correspondence of David Ricardo*, ed. P. Sraffa, Vol. 3, 1951. Cambridge: Cambridge University Press.
Silberling, N.J. 1924. Financial and monetary policy of Great Britain during the Napoleonic Wars. *Quarterly Journal of Economics* 38: 214–233.
Viner, J. 1937. *Studies in the theory of international trade*. New York: Harper & Row.

CHAPTER 21

Dollar-Sterling Exchange Market

Originally published in *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*, pp. 230–240, 267–276, 309–313, copyright Cambridge University Press, 1996. Reprinted with permission.

21.1 Empirical Testing of Market Efficiency

21.1.1 Unit of Observation

A perfectly efficient operation involves the pertinent exchange rate always within the efficiency band of that operation. Perfectly efficient gold-point arbitrage (GPA) requires absolute confidence in the maintenance of the gold standard over the time needed for the operation to be completed, specifically, in being able to transact in gold with both authorities. In turn, perfectly efficient uncovered interest arbitrage (UIA) and forward speculation (FS) (and in fact the applicability of the UIA, FS model) require full confidence in perfectly efficient GPA over the agent's horizon. Covered interest arbitrage (CIA) has no such condition for its efficiency, because that operation behaves the same way under a gold standard as under a flexible exchange-rate system.

For perfect efficiency, (1) perfect knowledge of the relevant parameters in the efficiency band (or, equivalently, in the associated profit formulas) and (2) instantaneous adjustment are also needed. Otherwise, only what may be called general efficiency—but not perfect efficiency—is attainable.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_21

General efficiency incorporates "inefficiencies" (violations of efficiency points) that are of small magnitude and non-persistent.

To test for the market efficiency of GPA, UIA, CIA, and FS, monthly data are used. So efficiency is tested only in the general sense. This means that measured efficiency is consistent with inefficiencies in the form of intra-month profit opportunities that average out for the month. While, for a given operation (GPA, UIA, CIA, FS), violations of efficiency points for isolated months indicate some inefficiency (even in the general sense), especially serious would be their persistence over several months.

21.1.2 Sample Periods

The interwar dollar-sterling gold standard provides the basic sample period: May 1925–August 1931, the 76 full months for which the standard applied. This experience is selected because it alone has the data and information to test all the operations for market efficiency. For comparison purposes, two other sample periods are adopted, one from the pre-World War I gold standard, the other from the post-World War II Bretton Woods system (gold-exchange standard).

The Bretton Woods experience is of interest because it possessed the two defining characteristics of a gold standard specified in Officer (1996, Sect. 2 of chapter 12). Currencies (except the US dollar) had a declared par value, quite analogous to mint parity. There was a band of 1% around the par value within which the member government—acting through an agency, such as its Treasury or central bank—was required to maintain spot exchange-rate transactions. This "parity band" corresponds to the gold-point spread. Indeed, the term "gold points" was used synonymously for the parity points or limits.

Britain was on the Bretton Woods system continuously from December 1946 to August 1971. The pound's par value was altered, however, by the depreciations of September 1949 and November 1967. So the maximum period encompassing full years of an unchanged par value is the seventeen-year period from 1950 to 1966. This is the period outside the span of the gold or specie standard, that is examined in the context of exchange-market integration in Officer (1996, chapters 10 and 11). It will play a similar comparison role under the rubric of GPA in this chapter.

For the pre-World War I era, a sample period of the same length is appropriate, and 1890 to 1906 is selected. This period is close to 1890–1908, which is adopted by Clark (1984) for his study of GPA.

The recurring lack of confidence in Britain‘s maintenance of the pound’s current par value during the Bretton Woods system is ably documented.[1] Depreciations comparable in magnitude to those of 1949 and 1967 could have occurred without surprise on several occasions between 1950 and 1966, because changes in the par value were institutionalized in the Articles of Agreement of the International Monetary Fund and because of Britain’s repeated foreign-exchange crises. During periods of sterling crisis—parts of 1951–1952, 1955–1956, 1957, 1961, and 1964–1965, involving eight years of the seventeen-year period—speculators did not necessarily have confidence in the maintenance of the pound’s par value and hence the parity band, and therefore could not be expected to provide the stabilizing speculation of UIA. Indeed, they probably engaged in the opposite behavior (destabilizing speculation).

One implication is that the UIA and FS models cannot be applied to the 1950–1966 period. These exclusions make the CIA model uninteresting, leaving only the analogy of GPA to be investigated. Another implication is even more serious. Does the tendency to destabilizing speculation as a characteristic of the Bretton Woods system bias the results for the remaining operation, GPA (or GPA analogue), toward market inefficiency compared to the 1890–1906 period?

The answer is negative, because during the eighteen-year period from 1890 to 1907 there was, similarly, an acute lack of confidence in the ability of the United States to remain on the gold standard, leading to similar episodes of destabilizing speculation. So again the UIA and FS models are inapplicable. There were recurrent runs on banks and bank failures, and for a one-month period in 1893 banks refused to provide cash for their deposits or even to honor certified checks (see Officer 1996, Sect. 9 of chapter 3). Coupled with the runs on banks were runs on the Treasury’s gold reserves. A $100-million gold reserve had a profound psychological impact as the “apprehension minimum” (quoting Bagehot) as well as a “legal minimum.”[2] At the end of 1893 the reserve was below $70 million and by February 1895 it had fallen to $41 million. The gold standard was on the verge of collapse at that point and was saved only by the cooperative action of the Treasury and a bankers’ syndicate.

Both contemporary observers and later historians agree that there was a fluctuating, and at times acute, lack of confidence in US maintenance of the gold standard on the part of both Americans and foreigners in the early and mid-1890s.[3] In respect to lack of confidence, the British Bretton Woods experience is strikingly comparable:

> The US experience in 1891-1897 was similar to British post-World War II exchange crises. In both cases, a government was seeking to maintain a fixed exchange parity. In both cases, it was uncertain whether the government would succeed. In both cases, it was clear that, if there were any change, it would be a depreciation of the relevant currency. Hence, in both cases, there was an incentive to reduce the balances held of the currency in question, and this incentive varied in intensity as the chances of the maintenance of fixed parity varied. (Friedman and Schwartz 1963, p. 105)

Of course, the actions of a banking syndicate in "saving" the gold standard in 1895 (and again in 1896, when gold exports were blocked) have no analogue under the Bretton Woods System.[4] Rather, the similarity is that exchange-rate speculation occurred against a background of periodic expectations of a change in the dollar/sterling parity for about the same aggregate time during each sample period. Therefore there is no clear sample bias of exchange-rate expectations on the empirical results to come.

It is also arguable that there could not possibly be gold-point violations in 1950–1966, as the United Kingdom (through its agent, the Bank of England) was under the international legal obligation to maintain the exchange rate within 1% of parity. However, the "gold points" of interest are not the legal maxima of 1% but rather the inner points, that were the announced intervention limits of the Bank of England and that were well under 1% (see Officer 1996, Sect. 3 of chapter 9).

21.1.3 Gold-Point Arbitrage

21.1.3.1 Conventional Wisdom and Empirical Studies

The quick working of gold-point arbitrage under the gold standard in the late nineteenth and early twentieth centuries, especially for the dollar-sterling foreign-exchange market, has long been part of the conventional wisdom of economic historians. The greatest historian of this market, referring to the period 1897–1913, wrote: "the foreign-exchange market in New York was given a steadiness theretofore unknown…Any appreciable deviation of exchange rates from parity tended to provoke a movement of gold into and out of the country" (Cole 1929, p. 213). The foremost historian of foreign-exchange in general, stated: "But on the whole the behavior of the sterling-dollar rate conformed to the 'rules of the game'and kept within gold points for some thirty-five years prior to 1914" (Einzig 1970, p. 195).

Contradicting this conventional wisdom, the empirical studies of Morgenstern (1959) and Clark (1984), supported by the incidental finding of Moggridge (1972), call into question the celebrated GPA efficiency of the US-UK gold standard. Coining the term "violations of gold points" for observations of the exchange rate outside the gold-point spread and hence profitable opportunities for GPA, Morgenstern (1959, pp. 246–248, 252–253) finds 45 months of such violations in 1880–1914 and eleven months in 1925–1931. Moggridge (1972, pp. 128, 185) discovers ten months of GPA inefficiency in 1925–1931. The results of Clark (1984, pp. 800–803) are strongest, with GPA inefficiency occurring in 47 months in 1890–1898 and 51 weeks in 1899–1908—a total of 98 instances. Perhaps most disturbing to believers in market efficiency is that the gold-point violations are often consecutive over several periods.

These authors have cause to trumpet their findings. Morgenstern (1959, p. 276) highlights "the incredible phenomenon of exchange rates often and persistently beyond the gold points." Clark (1984, p. 818) comments: "It was found that gold point violations often persisted for several successive months. This suggests that the gold standard system did not always eliminate profit opportunities quickly…These findings are inconsistent with the view that the gold standard system functioned efficiently."

21.1.3.2 Empirical Testing

Data and Results

Only tests of GPA efficiency are presented here. For testing UIA, CIA, FS, see Officer (1996, pp. 240–51). To test for GPA efficiency—the exchange rate between the gold points, $GM \leq RS \leq GX$ or, equivalently, $-CM \leq S \leq CX$—monthly values are needed for the spot exchange rate (RS or S) and gold points (GX and GM, or CX and $-$ CM) for 1890–1906, 1925–1931, and 1950–1966. It is convenient to consider the variables in the form of percentage deviation from parity: S, $-$CM, and CX. Definitions of all these variables are in Sect. 23.2.2. Their time series are developed in Officer (1996, chapters 6 and 9).

GPA inefficiencies are listed in part I of Table 21.1. The charge that the gold standard at the turn of the century was beset by GPA inefficiencies is found to be false. There are only three, isolated, months of a gold-point violation in 1890–1906. Also, in the interwar period GPA was

Table 21.1 Gold-point violations: 1890–1906, 1925–1931, 1950–1966

Year	*Month*	*Exchange rate* (Percentage deviation from parity)	*Gold Point* (Percentage deviation from parity)	*Magnitude of violation*[a] *(Percent)*
Part I: Gold-point Arbitrage				
Violations of Gold-Import Point				
1891	October	−0.6898	−0.6680	3.27
1893	August	−0.1258	0.3447	136.50
1906	September	−0.6476	−0.6445	0.48
1951	September	−0.0464	−0.0446	4.01
Part II: Transfer of Funds				
Violations of Gold-Import Point				
1890	March	−0.2673	−0.2555	4.61
	September	−0.4109	−0.2555	60.81
	December	−0.5381	−0.3918	37.34
1891	September	−0.5852	−0.3095	89.10
	October	−0.6898	−0.3092	123.07
	November	−0.6283	−0.3091	103.29
	December	−0.5175	−0.3090	67.50
1892	January	−0.3330	−0.3087	7.85
1893	July	−0.5750	−0.3597	59.86
	August	−0.1258	0.7751	116.23
	October	−0.4642	−0.3591	29.27
	November	−0.4129	−0.3589	15.05
1896	September	−0.5695	−0.3889	46.43
	October	−0.5140	−0.3619	42.01
1897	October	−0.4051	−0.3928	3.14
1898	January	−0.3948	−0.3664	7.76

Year	*Month*	*Exchange rate* (Percentage deviation from parity)	*Gold Point* (Percentage deviation from parity)	*Magnitude of violation*[a] *(Percent)*
	March	−0.5839	−0.3669	59.13
	April	−0.6352	−0.3672	72.99
	September	−0.5181	−0.3687	40.51
	October	−0.4236	−0.3691	14.77
	December	−0.4565	−0.3696	23.48
1899	January	−0.4236	−0.3700	14.49
1900	October	−0.4996	−0.3968	25.91
	November	−0.4174	−0.3971	5.12
1903	November	−0.5900	−0.3327	77.35
	December	−0.5962	−0.3330	79.04
1906	April	−0.4133	−0.0408	913.20
	May	−0.3393	−0.1231	175.67
	July	−0.4010	−0.3357	19.45
	August	−0.4195	−0.3766	11.39
	September	−0.6476	−0.0577	1022.68
	October:	−0.3743	−0.1808	107.01
	December	−0.6208	−0.3902	59.12
1925	October	−0.4947	−0.4384	12.85
1951	September	−0.0464	−0.0446	4.01
Violations of Gold-Export Point				
1895	August	0.6351	0.5416	17.26
	December	0.4833	0.4543	6.37
1896	January	0.4641	0.4543	2.15

[a]Deviation of exchange rate from gold point as percentage of gold point, absolute value
Source See text

uniformly efficient on a monthly average basis. For comparison, there was one violation of a "gold point" in 1950–1966.

The magnitude of a gold-point violation can reasonably be defined as the deviation of the exchange rate from the violated gold point as a percentage of the gold point: $100 \cdot [(S + CM)/(-CM)]$ and $100 \cdot [(S - CX)/CX]$ for an import and export-point violation, respectively. All four GPA inefficiencies are with respect to the US gold-import point. Only for August 1893 is the inefficiency large.

It might be mentioned that the taxation element turns out to be irrelevant for GPA efficiency, and not only because of its proportional effect on GPA profit. While the US had a corporate income tax in 1925–1931, there were no gold-point violations during that period. The gold-point violations in 1890–1906 occurred when there was no US income tax.[5]

Explanations of Gold-Point Violations

The differential pattern of GPA inefficiency in the Morgenstern-Moggridge-Clark studies versus the present findings requires explanation. Also, there must be an accounting of the (albeit only four) gold-point violations discovered here. Any test of market inefficiency is a joint test of several hypotheses. In particular, the hypotheses are (1) the specified values of the variables and parameters of the model of GPA equilibrium, (2) the validity of the model itself, and (3) GPA efficiency. Each hypothesis is considered in turn, with applicability as warranted to the three existing and the current findings of gold-point violations.

Values of Variables and Parameters

Exchange rate: Morgenstern, Moggridge, and Clark employ the cable-transfer exchange rate throughout their time periods, with the exception of Morgenstern for 1880 to August 1886. In the interwar period cable was certainly the medium for GPA, but decidedly the bill of exchange was dominant prior to World War I, and in the form of the demand bill from about 1880 (see Officer 1996, Sect. 5 of chapter 8). In adopting the cable transfer for the pre-World War I period, Morgenstern and Clark dealt with an instrument that typically was eschewed by gold-point arbitrageurs.

The way the cable exchange rate may have led to greater measured inefficiency than if the, correct, demand-bill rate had been used is as follows. The cable rate exceeds the bill rate, because the latter incorporates an interest component emanating from the duration of a New York to London Atlantic voyage (see Officer 1996, Sect. 2 of chapter 6).

The cable rate, being too high, could lead to measured inefficiencies under gold-export arbitrage. It is also true that it could enhance measured efficiency under gold-import arbitrage.

Gold points: There are several reasons why the gold points of Morgenstern, Moggridge, and Clark are probably narrower than the true gold points, leading to spurious cases of GPA inefficiency. First, unlike the present writer, these authors do not explicitly respect the difference between GPA and gold-effected transfer of funds [GTF] gold points (discussed in Officer 1996, chapter 8). Except for Clark's development of the interest-cost component of arbitrage cost, none of the previous authors (Morgenstern, Moggridge, and Clark) generates gold-point data from first principles (see Officer 1996, Sect. 1 of chapter 9). The outcome is uncertainty as to whether the gold points that they use pertain to GPA, GTF, or some mix.

The issue is relevant because GTF points are inherently smaller in magnitude (see Officer 1996, Sect. 3 of chapter 9). To the extent that the gold-point data utilized apply to GTF rather than GPA, the gold-point violations of the three authors may be spurious. To illustrate the quantitative impact of this element, consider part II of Table 21.1, which exhibits inefficiencies for GTF. The three gold-point violations under GPA in 1890–1906 become 33 under GTF!

Second, Clark (1984, pp. 798–799) substitutes the product of the spot exchange rate and the US/UK interest-cost ratio for the forward exchange rate in his export-arbitrage profit formulas. This substitution ignores transactions costs in the countries' money markets, a deficiency recognized by Clark (1984, p. 804) as overestimating his arbitrage profitability figures, that is, GPA inefficiencies.

Third, the gold points of all three authors, like most estimates extant, are probably unduly small in magnitude, because of the omission of direct mint expenses, interest forgone in waiting time for mint procedures in selling gold, exchange-rate cost (recognized only in principle by Clark), risk premium, and normal profit. The use of existing estimates rather than construction of gold points from first principles, the latter done only by the present author, gives scope for such omissions.

Fourth, the technique of imposing invariant gold points over time, utilized by the three authors, ignores the trend of smaller gold points over time, especially for 1890–1906. When Clark allows for a higher cost in the early part of the period, the number of GPA inefficiencies drops from 98 to 23.

Validity of Model of Equilibrium

Assumption of confidence in the gold standard: If gold-point arbitrageurs lack confidence that the gold standard will be maintained by one of the countries, then the model of GPA equilibrium—involving gold flows (disequilibrium) when the exchange rate is outside the gold-point spread—may become inapplicable. During the 1890s, the view that the United States might abandon the gold standard is well documented (see Sect. 21.2). How would this belief affect the behavior of gold-point arbitrageurs? These agents would have had a positive subjective probability of a dollar depreciation in terms of gold and the pound, either because the US mint price would be increased or because the US would go on a paper standard. The resulting incentive would be to import and not to export gold. However, not to engage in gold-export arbitrage involves the surrender of an immediate profitable opportunity.

What is the evidence regarding arbitrageur's balancing of these considerations? The present study finds no export-point violation during 1891–1896, and only two months of import-point violation. It appears that arbitrageurs had a sufficiently short time horizon and/or a sufficiently low probability of a dollar depreciation in the near term, that they typically took advantage of GPA profit opportunities, whatever the direction of the gold-point violation.

Assumption of authorities transacting in gold with private parties: In 1950–1966 the US and UK authorities did not pursue a policy of transacting in gold in both directions with all comers. The countries were not on the gold standard, and GPA could not be consummated. The so-called GPA inefficiency in September 1951 reflects the Bank of England not respecting its self-imposed lower intervention point, allowing the pound to depreciate against the dollar beyond the announced limit (but still well within the 1-percent parity band mandated by the International Monetary Fund).

Efficiency

Intra-monthly observations: Clark's exchange-rate observations pertain to specific days (one per month or week) and the interest-rate component of his gold-point figures is computed from weekly interest-rate data. In contrast, the exchange-rate and gold-point data of the present study are (actually or essentially) monthly averages of daily observations. It is possible that some of the inefficiencies found by Clark are genuine but of the intra-month variety, that would be eliminated on a monthly-average basis. However, this explanation is of small importance, because the preponderance of Clark's gold-point violations is consecutive for two or more time periods.

Gold-point manipulation: Clark's principal explanation of gold-standard inefficiencies in 1890–1908 is direct manipulation of gold points by the Bank of England and US Treasury. However, his discussion of the policies employed is incomplete, reveals an awareness of neither Sayers' (1936) definitive work on the Bank's gold-market operations nor the existence of the Treasury's bar premium, and offers no specific dates beyond the Treasury's well-known actions under Secretary Shaw in 1906 (Clark 1984, pp. 816–817). While the policies are discussed in detail in Chapters 17–18 and Officer (1996, chapter 9), an outline of their effects on the gold points follows.

Bank actions that lowered the American gold-export point, thereby enhancing a gold inflow into Britain, consisted of raising its purchase prices for bars and foreign coin above normal levels and making interest-free loans to gold importers. Bank policies that lowered the American gold-import point, discouraging gold outflow from Britain, involved raising selling prices for bars and foreign coin above normal levels, refusing to sell bars, and redeeming its notes in deliberately underweight sovereigns.

Treasury actions that raised the gold-export point, discouraging a gold outflow, were raising the bar premium above the normal rate (0.04%—see Officer 1996, Table 9.5) and refusing to sell bars outright. The Treasury's policy of refusing to provide double-eagles exclusively but paying out coin in proportion to the various denominations in stock cannot logically be considered a gold-point manipulation. Finally, in making interest-free loans to gold importers, the Treasury raised the gold-import point, fostering a gold inflow.

Clark (1984, pp. 791, 814–819) gives as the reason for the Bank and Treasury gold-point manipulation their desire to delay the workings of the gold standard insofar as it imposed limits on discretionary monetary policy, in particular, the intent to circumvent the obligation to reduce the money supply when gold reserves fell.[6] Sayers (1936, pp. 71–101), in contrast, developed a full history of the Bank's gold-point manipulation based on the view that changing the gold points was an alternative to increasing the Bank's rediscount rate (Bank Rate) when imbalance was external and it was desired not to upset internal balance to stem a gold outflow or encourage an inflow. Also, the Bank's gold-point intervention complemented a Bank Rate change when the latter was ineffective either in the sense of not governing the market rate or in the sense of not affecting gold flows sufficiently.[7]

The difference between the Clark and Sayers interpretations is substantive, even though both view gold-point manipulation as an alternative to monetary policy. For Clark, that monetary policy is passive, the money supply changing with gold reserves in accordance with the "rules of the game" of the gold standard. For Sayers, it is active, with gold-market policy used in place of, or in conjunction with, Bank-rate policy aggressively designed to foster a net capital inflow and improve gold reserves. Sayers' explanation is supported by (1) a detailed description and analysis of the Bank's interest-rate and gold-market policies in the 1890–1914 period and (2) reference to contemporary newspaper reports and commentaries that stated, and encouraged, the relationship between Bank-rate and gold-point policies. In contrast, Clark's argument for his position is founded entirely on his incorrect cost estimates of shipping gold.

The Treasury's motivations for gold-point operations were quite different from those of the Bank, but again were incorrectly described by Clark. Secretary Shaw's interest-free advances in 1906 were unrelated to the state of the Treasury's gold reserves; the objective was to provide extra funds for commercial banks, thereby easing seasonal or cyclical pressure in the money market.[8] As for the Treasury's bar embargo and high premium in the 1890s, these were part of "a sequence of short-sighted expedients to obtain and retain gold" in the face of a lack of confidence in US maintenance of the gold standard and "a series of flights and returns [from/to the dollar] as views altered" (Friedman and Schwartz 1963, pp. 106, 104).

Irrespective of the motivation for gold-point manipulation, examination is required of Clark's judgment that such policies constituted "interference in the workings of the gold standard system," even stating that the Bank and Treasury "tampered with the free workings of the gold market" (Clark 1984, pp. 792, 818).[9] On the contrary, the Bank and Treasury always adhered to the basic statutes that placed their respective countries on the gold standard. For the Bank, the Resumption Act of 1819 required it to cash its notes in sovereigns at the mint price, while the Bank Charter Act of 1844 mandated that it purchase gold bars at 77s. 9d. per standard ounce (see Sects. 18.4 and 18.5). For the Treasury, the

Resumption Act of 1875 stipulated gratuitous coinage of standard gold bullion (meaning that only the expenses of converting bars and foreign coin into US coin would be assessed and not the additional charge of 0.2% specified in the Mint Act of 1873). Also, the Treasury was required to redeem US notes, Treasury gold certificates, and Treasury notes in coin by the acts of 1875, 1882, and 1890 (see Sects. 17.4 and 17.5).

These obligations, always followed in the 1890–1914 period, placed outer limits on the dollar-sterling gold points, given (1) freight, insurance, and handling costs, (2) other transactions costs (abrasion, mint charges, currency premium, and exchange-rate cost), (3) interest cost, and (4) normal profit and risk premium. In offering to buy or sell foreign coin and to sell bars, at any price for which there were takers, and in making interest-free advances, the Bank went beyond statutory requirements and narrowed the gold-point spread, thus enhancing external integration but in no way impairing GPA efficiency. In supplying bars, at any premium for which takers existed, and in making interest-free advances, the Treasury did the same.

As the gold-market policies of the Bank and Treasury varied over time, the gold-point spread would narrow or widen, but it would never go beyond the outer limits emanating from obligations under the basic statutes. Arbitrageurs responded to the altered cost of receiving or selling gold as they would to a change in transport charge, insurance rate, interest cost, or any other component of arbitrage cost. There was no hindrance to GPA efficiency. In sum, gold-point manipulation did nothing more than alter the costs of gold shipments, but with the costs always remaining less than in the absence of such manipulation and in the presence of statutory observance by the Bank of England and the US Treasury.

Banking syndicates: Clark (1984, p. 817) puts forward another cause of gold-standard inefficiencies: the activities of banking syndicates in preventing gold exports. He discusses two such syndicates, the Morgan-Belmont group, operating in February–September 1895, and a similar syndicate functioning in July–August 1896.[10] Clark's results show that all ten months exhibit gold-export point violation.

To the extent that these syndicates succeeded in putting pressure on gold exporters to act against their economic interest via threats or bribes, then gold-point violations could occur. However, the present study shows no GPA inefficiencies during the periods when the syndicates functioned (and in fact not one gold-export violation on a monthly average basis during 1890–1906). So while in principle banking syndicates could have induced arbitrageurs to desist from profitable activity, the evidence is that they did not have this effect in a measurable monthly way.

Need for liquidity: When New York bankers, who were the gold-point arbitrageurs, desired quick liquidity, they could sell sterling bills for immediate domestic currency, and were especially likely to do so during a financial crisis. The increased supply of bills could push the exchange rate below the gold-import point. There was no immediate correction via GTF or GPA, because of the delay in obtaining funds via gold import. Ironically, the gold-point arbitrageurs (the New York banks) bring about the gold-point violation, rather than correct an exogenously imposed one (from their standpoint).[11] This is almost certainly the explanation of the only substantial GPA inefficiency in Table 21.1, that of August 1893. The financial crisis of 1893 was most severe in that month, when a premium on currency occurred (see Sect. 17.9).

21.2 Net Outcome

21.2.1 *Methodology of Measuring Regime Efficiency*

21.2.1.1 Determinants of Regime Efficiency

The efficiency bands of the various operations—gold-point arbitrage, uncovered interest arbitrage, covered interest arbitrage, and joint "forward speculation with covered interest arbitrage"—the extent to which the operations are efficient (within the respective bands), and the influences of the various market and policy variables combine with the determinants of external and internal integration to generate a certain level of regime efficiency. So interest now extends beyond 1925–1931 to 1791–1931, 1950–1966—the entire time span of the study (Officer 1996). What is the appropriate measure of the extent of regime efficiency?

21.2.1.2 A Simple Measure of Efficiency

Perfect regime efficiency is defined as the exchange rate at the midpoint of the gold-point spread. At this point it is convenient to revert to expressing the exchange rate in percentage points of parity and as the deviation from the spread midpoint: R^*, as developed in Officer (1996, Sect. 1 of chapter 11). So perfect regime efficiency is given by $R^* = 0$. A simple measure of the amount of regime inefficiency is then $|R^*|$. This is the experienced (observed) deviation of the exchange rate from the spread midpoint. For a given period of time with a time series of $|R^*|$, the average of the deviations may be taken; so the measure of inefficiency becomes mean$|R^*|$.

The lower bound of mean$|R^*|$ is zero, for the exchange rate (in dollars per pound) uniformly at the spread midpoint (RS at RM), where there is perfect regime efficiency. There is not an unconstrained upper bound. However, under perfect GPA, there is an upper bound (minimum regime efficiency or maximum inefficiency) of half the gold-point spread in percentage terms, given by CX*, the gold-export point expressed in percentage points of parity and as the deviation from the spread midpoint (the notation again from Officer 1996, Sect. 1 of chapter 11). The upper bound is reached for the exchange rate at either gold point (RS at GM or GX, in dollars per pound), where there is maximum regime inefficiency given efficient GPA.

There are several problems with mean$|R^*|$ as the measure of regime efficiency. First, it provides a biased representation of regime efficiency, because there is likely to be a positive correlation between mean$|R^*|$ and the width of the gold-point spread. After all, under efficient GPA the gold-point spread constrains the exchange rate to be within it. Second, the measure is not explicitly derived from a loss function. Third, it is not developed as a special case of a class of measures emanating from a general loss function. Fourth, there is no standard of attainable (rather than perfect) regime efficiency, comparable to the concept of general market efficiency. An explicit modelling approach to measuring regime efficiency corrects these problems.

21.2.1.3 A Model of Measurement of Efficiency

Loss Function

Logical properties of the loss function relating the disutility from regime inefficiency to the deviation of the exchange rate from the spread midpoint are as follows. (1) Disutility is zero at the spread midpoint. (2) At other exchange rates, disutility is positive. (3) The loss increases with the distance of the exchange rate from the spread midpoint. (4) The loss depends on only the absolute deviation of the exchange rate from the midpoint and not whether the deviation is positive or negative.

Two fundamental loss functions with these properties are the identity function (disutility equals deviation itself) and square function (disutility equals square of deviation). The square function especially penalizes exchange-rate deviations the further they are from the spread midpoint. Disutility increases at a higher rate than for the identity function.[12]

Experienced Versus Hypothetical Loss

Experienced loss: Given any sample period of a gold (in general, specie) standard, the deviation of the exchange rate from the spread midpoint—this deviation expressed as a percentage of parity, taken as positive, and denoted as $|R^*|$—can be constructed for all observations. Then the experienced average loss from regime inefficiency is the average deviation, mean$|R^*|$, or the average of the squared deviations, mean$(|R^*|^2)$, depending on the specific loss function. These are the observed average disutilities relative to the (zero) loss under perfect regime efficiency ($R^* = 0$ for all observations).[13]

Hypothetical loss: Consider the case of R^* a continuous random variable with density (probability distribution function) f(R^*). Assuming perfect GPA efficiency, f(R^*) = 0 for $|R^*| > \mathrm{CX}^*$, where CX^* is the gold-export point (percent of parity) expressed as the deviation from the spread midpoint. With positive GPA cost, $\mathrm{CX}^* > 0$. Assume that UIA and CIA are "neutral" in the sense that, combined with non-arbitrage and non-speculation market forces, they distribute the exchange rate along all points within the spread with equal probability. Then (from Officer 1996, Sect. 1 of chapter 11) $-\mathrm{CM}^*$ denotes the gold-import point, and $\mathrm{CX}^* = \mathrm{CM}^*$, whence R^* has the uniform (rectangular) distribution:

$$\begin{aligned} \mathrm{f}(R^*) &= 1/\big(\mathrm{CX}^* - (-\mathrm{CM}^*)\big) = 1/\big(\mathrm{CX}^* + \mathrm{CM}^*\big) \\ &= 1/\big(2 \cdot \ \mathrm{CX}^*\big) \text{ for } \left|R^*\right| \le \mathrm{CX}^* \\ &= 0 \text{ for } \left|R^*\right| > \mathrm{CX}^* \end{aligned}$$

Hypothetical average disutility emanates from this distribution, and depends only (and positively) on the gold point, CX^*. Therefore average disutility is positively correlated with the width of the gold-point spread. The hypothetical average loss from regime inefficiency is again the average exchange-rate deviation or average squared exchange-rate deviation (depending on the specific loss function), but the exchange-rate function follows the assigned uniform distribution. The deviation is on average half the magnitude of a gold point $\big(\mathrm{CX}^*/2\big)$, and this is the average hypothetical disutility under the identity loss function. For the square loss function, average hypothetical disutility is $\big(\mathrm{CX}^*\big)^2/3$, though intuition must give way to mathematics for the derivation.[14]

Efficiency ratio: With both the actual and hypothetical average losses measured relative to the zero disutility of perfect regime efficiency (and, of course, calculated for the same sample period), the ratio of the average actual to the average hypothetical disutility measures actual loss from regime inefficiency relative to the loss emanating from a hypothetical situation of perfect GPA and a uniform exchange-rate distribution within the gold-point spread.[15] Multiplication by 100 expresses the ratio in percentage terms. Insofar as the ratio is below (above) 100, regime efficiency may be accepted (rejected) for the sample period. A ratio of 100 is the criterion of attainable regime efficiency.

The efficiency ratio is $200 \cdot [\text{mean}|R^*|/\text{CX}^*]$ for the identity loss function and $300 \cdot [\text{mean}(|R^*|^2)/(\text{CX}^*)^2]$ for the square function.[16]

The lower bound of the efficiency ratio is zero, which occurs under perfect efficiency ($R^* = 0$ for all observations): the exchange rate is always at the spread midpoint. Under perfect GPA, the upper bound (minimum regime efficiency or maximum inefficiency) is reached when the exchange rate is always at a gold point ($|R^*| = \text{CX}^*$). Then, $\text{mean}|R^*| = \text{CX}^*, \text{mean}|R^*|^2 = (\text{CX}^*)^2$; the efficiency ratio is 200 for the identity loss function and 300 for the square function.

Relaxation of Assumptions

Uniform distribution function: The analysis could proceed just as well with the exchange rate having a hypothetical non-uniform distribution about the gold-point spread. If the exchange rate has a greater probability the closer it is to the spread midpoint, then the efficiency standard becomes tougher; whereas if the rate has higher probability as it approaches a gold point, then the standard is easier. At one extreme, hypothesized perfect efficiency places the exchange rate always at the midpoint, and efficiency for the given sample period can never be accepted. At the other extreme, perfect GPA but perverse behavior of other economic agents can be assumed such that the exchange rate is always at one or the other gold point. Now efficiency would be easy to accept. A uniform distribution appears to yield a good compromise standard.

Loss function: The model requires only that disutility be related to the deviation of the exchange rate from some norm value. Any of the four properties of the loss function can be dropped. The function can be discontinuous. For example, disutility can be made zero for the exchange rate within the spread and positive outside the spread. The function can be asymmetrical, as would occur, for example, by substituting mint parity for the spread midpoint as the norm value of the exchange rate.

Exchange-rate system: The model can be applied to any exchange-rate system involving a band of floating-rate activity with the band delimited by some effective force. For the gold standard, that force was private arbitrage and speculation, underlain by the willingness and ability of domestic authorities to transact in gold with the agents of these operations. Under the Bretton Woods system, the force was exchange-rate setting by domestic authorities. The model is equally applicable to either system.

Relationship to Internal Integration

On the surface, the regime-efficiency ratio appears to be simply another way of measuring internal integration. By definition, with II denoting internal integration and EI external integration (see Sect. 23.2.3), II = mean$|R^*|$ − EI/2, where EI = CX*; so II = mean$|R^*| - CX^*/2$. This formula may be compared to the efficiency ratio for the identity loss function: $200 \cdot (\text{mean}|R^*|/CX^*)$. Apart from the constants, the only difference is that internal integration is of the form A − B and the efficiency ratio the form A/B. Is this an essential difference and do the concepts differ in any other respects?

In fact, internal integration and regime efficiency, though related, are distinct in both concept and measurement. First, internal integration pertains to the perfection of the foreign-exchange market, whereas regime efficiency is concerned with the stability of the monetary standard. Second, the statistic mean$|R^*|$ is in part an indicator of *variation* of the exchange rate under exchange-market integration, gauging temporal as well as placement integration (see Officer 1996, Sect. 3.1 of chapter 11), but is purely a measure of *central tendency* for regime efficiency. Third, internal integration was developed under a specific (linear) formula, while regime efficiency is consistent not only with nonlinear formulas but also with entire families of alternative measures.

Fourth, regime efficiency is a stricter concept than internal integration. Stability of the monetary system (say, a gold standard) might be dependent strongly on the location of the exchange rate. Efficient GPA definitely keeps the exchange rate within its (the GPA) gold-point spread; the same cannot be said for "efficient" GTF in relation to its spread. GTF only plays a role in preventing a gold-point violation from worsening in frequency and magnitude (see Officer 1996, Sect. 6 of chapter 8). Therefore only the GPA spread and midpoint are used to construct R^* and CX*, the ingredients in the numerator and denominator of the regime-efficiency ratio. In contrast, internal and external integration were legitimately computed alternatively under GPA and GTF (see Officer 1996, Table 11.1). External integration is a dimension of market integration but not of regime stability; so the role of GTF can be considered alternative to GPA only for integration.

21.2.2 Empirical Results

Table 21.2 presents the average experienced exchange-rate deviation from the spread midpoint, the average hypothetical deviation, and the corresponding efficiency ratio for both the identity and square loss functions by period. The statistics are calculated for two sets of periods: (1) the usual periods spanning 1791–1800 to 1950–1966 and based on quarterly exchange-rate observations, (2) the periods for which monthly exchange-rate observations have been developed (1890–1906, 1925–1931, 1950–1966). The figures in Table 21.2 pertain entirely to the GPA gold-point spread, and it is no coincidence that the average experienced and hypothetical deviations for the identity function are identical (except for fewer significant digits) to the exchange rate and half the gold-export point in Officer (1996, Table 11.1) (columns two and four in each table). Figures 21.1, 21.2, and 21.3 plot gold points, spread midpoint or parity, and the exchange rate monthly for 1890–1906, 1925–1931, and 1950–1966.

Considering first the results based on quarterly observations, the tremendous decline in the average experienced exchange-rate deviation in the 1820s compared to the 1790s is even more apparent with the square than the identity function, because of the existence of extreme observations in the 1790s. In the 1860s, the decade incorporating the Civil War, the identity function shows average deviation hardly above that of the 1850s, but the square function exhibits a substantial increase. Thereafter, to World War I, the average deviation of the exchange rate from the spread midpoint falls steadily to the amazingly low figures of 0.12 and 0.02% of parity under the identity and square functions, respectively. In 1925–1931 average deviation increases to levels intermediate between those of the 1880s and 1890s. The deviations for 1950–1966 are about the same as for the interwar period.

Dividing the average experienced deviation by the average hypothetical deviation and multiplying by 100, the efficiency ratio is obtained. Results are most interesting. Unlike for external and internal exchange-market

Table 21.2 Exchange-rate deviation from spread midpoint, gold-point arbitrage, 1791–1966

Period	*Mean experienced deviation*[a]		*Mean hypothetical deviation*[a]		*Efficiency ratio*[b]			
	Identity function	*Square function*	*Identity function*	*Square function*	*Identity function*		*Square function*	
Quarterly observations								
1791–1800	4.19	27.97	3.56	16.94	117.42		165.10	
1821–1830	2.14	6.33	1.98	5.22	108.37		121.24	
1831–1840	1.44	3.74	1.81	4.35	79.96		86.07	
1841–1850	1.04	1.54	1.26	2.12	82.59		72.88	
1851–1860	0.86	0.92	0.80	0.85	108.14		108.51	
1861–1870	0.87	1.52	0.74	0.74	116.85		205.40	
1871–1880	0.38	0.23	0.55	0.41	68.70		56.37	
1881–1890	0.32	0.15	0.34	0.16	93.46		97.74	
1891–1900	0.25	0.08	0.32	0.14	77.03		61.98	
1901–1910	0.15	0.04	0.27	0.10	53.34		35.15	
1911–1914[c]	0.12	0.02	0.27	0.10	42.66		20.59	
1925[d]–1931[e]	0.28	0.09	0.27	0.10	102.83		95.44	
1950–1966	0.26	0.10	0.32	0.14	80.48		75.57	
Monthly observations								
1890–1906	0.24	0.09	0.30	0.13	78.37	77.62[g]	71.88	64.90[g]
1925[e]–1931[f]	0.28	0.10	0.27	0.10	104.01	102.97[g]	97.24	94.21[g]
1950–1966	0.27	0.11	0.32	0.15	85.21	87.18[g]	76.38	78.16[g]

[a]Percentage points of parity
[b]Ratio of mean experienced to mean hypothetical deviation, percent, except where note g
[c]Second quarter
[d]Third quarter
[e]May
[f]August
[g]Mean of monthly ratios of experienced to hypothetical deviation

integration, the improvement in regime efficiency in the 1820s is by no means overwhelming. Indeed, for the identity function (comparable to II, the measure of internal integration), there is a greater decline in the ratio (and therefore increase in regime efficiency) from the 1820s to the 1830s than from the 1790s to the 1820s.

The criterion of attainable regime efficiency (a 100-percent ratio) is accomplished by the 1830s, lost in the 1850s and 1860s, but then regained uniformly except (almost but not quite) for the identity function

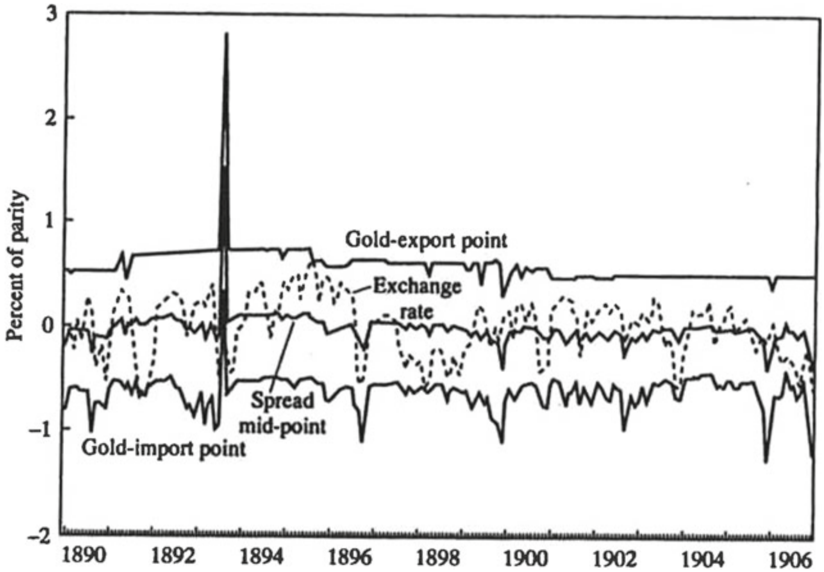

Fig. 21.1 Exchange rate and gold points, gold-point arbitrage, monthly, 1890–1906

in 1925–1931.[17] Again, in 1911–1914 regime efficiency is at a remarkably high level: 43% for the identity function and 21% for the square function. The regime-efficiency ratio suggests that World War I marked a watershed in gold-standard stability. Regime efficiency in the interwar period worsened noticeably. For the identity function, the efficiency ratio more than doubled, so that regime efficiency was no longer attained. For the square function, the ratio more than quadrupled. In contrast, in 1950–1966 regime efficiency is attained, with the ratios again between the levels for the 1880s and 1890s.

Turning to the monthly observations, they show that there is little precision forgone when quarterly observations are used to compute the average experienced deviation of the exchange rate. Values are about the same for corresponding time periods irrespective of unit of observation (month or quarter). Also, in five of six cases (the square function for 1890–1906 the only exception), the difference between computing the efficiency ratio as the ratio of the monthly average experienced to the

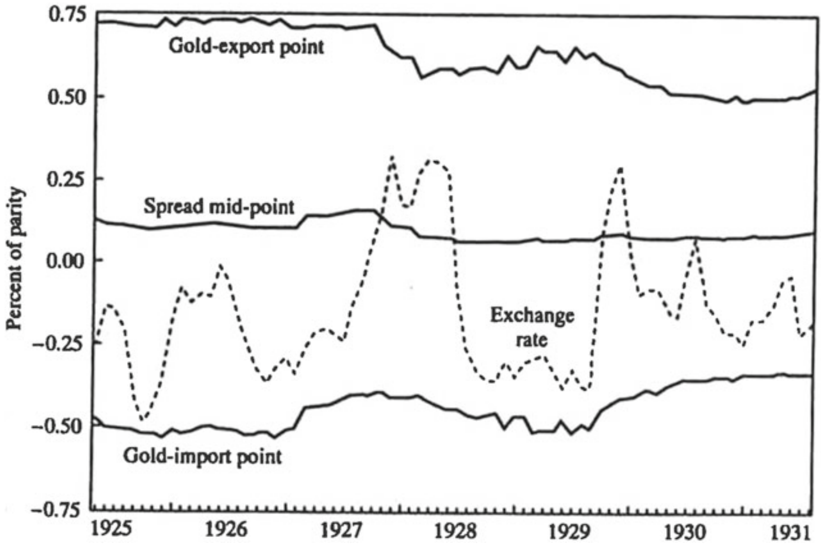

Fig. 21.2 Exchange rate and gold points, gold-point arbitrage, monthly, 1925–1931

monthly average hypothetical deviation or as the average of the monthly experienced/hypothetical ratios is fairly small.

Returning to the interwar period, regime efficiency is worse than for the other monthly periods, 1890–1906 and 1950–1966. Yet, whether monthly or quarterly exchange-rate observations are used, regime efficiency is accepted for the square loss function and barely rejected for the identity function. In either situation the ratio is in the neighborhood of 100. This means that regime efficiency in 1925–1931 was about the same as that emanating from efficient GPA and a uniform distribution of the exchange rate within the gold-point spread. This is perhaps a surprising result, in view of the long-run regime inefficiency of the interwar gold standard. However, long-run does not necessarily translate into short-run regime inefficiency (see Officer 1996, Sect. 1 of chapter 14), and the latter is what the efficiency ratio measures.

The important conclusion is that the short-run regime efficiency of the 1925–1931 dollar-sterling gold standard was far greater than the conventional wisdom concedes. In this the role of private market participants was positive if not paramount, as shown in the results of Officer (1996,

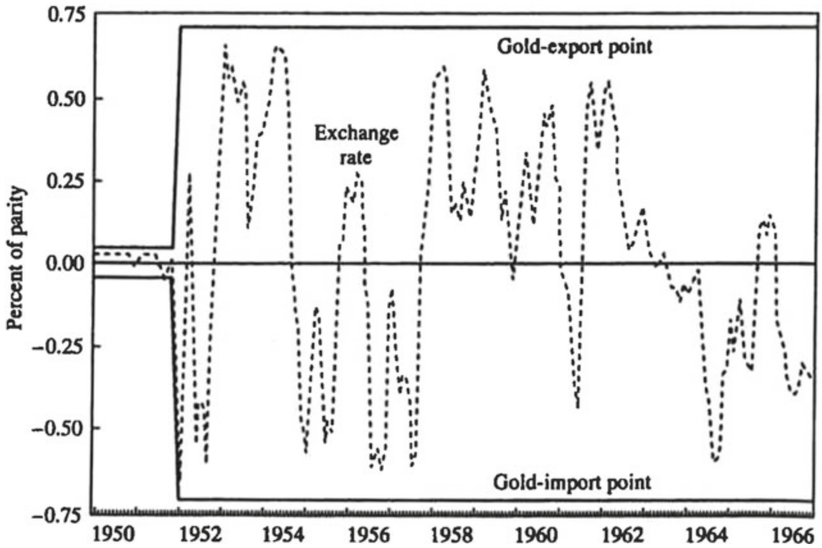

Fig. 21.3 Exchange rate and gold points, monthly, 1950–1966

chapters 13–14). Until confidence in the gold standard collapsed in July 1931, GPA exhibited complete market efficiency (on a monthly unit of observation). Further, from May 1925 through June 1931, speculators and arbitrageurs in the foreign-exchange and money markets behaved as if any lack of confidence that they had in the gold standard and in efficient GPA was episodic rather than systemic. Also, the likelihood of the expected exchange rate being close to the spread midpoint and the location of the UIA and CIA bands in relation to the gold-point spread enhanced regime efficiency.

However, as shown in Officer (1996, chapters 13–14), market forces were not uniformly conducive to regime efficiency. The extreme risk aversion of exchange-rate speculators and interest arbitrageurs and the non-negative correlation of the London–New York market interest-rate differential with the strength of the pound detracted from efficiency. Further, the interest-rate differential often had the sign opposite of that mandated for regime efficiency.

Short-run policy on the part of the monetary authorities of the United Kingdom and United States offered precious little assistance to regime efficiency, as found in Officer (1996, chapter 15). The international discount-rate differential (Bank Rate minus Federal Reserve Bank of New York rediscount rate) was generally too low to affect the market interest-rate differential significantly and further was not oriented to the weakness of the pound. Exchange-market intervention acted to weaken the pound in a large minority of months, and, when supportive, was of low magnitude on average.

On balance, it was the operations of private participants rather than of monetary authorities that gave regime efficiency the strength that it had in the interwar period.

Notes

1. For an excellent account, with extensive references to the literature, See Yeager (1976, pp. 441–472).
2. "So long as the Treasury's surplus gold fund has held above $100,000,000, the public mind has been generally easy; when the gold has fallen below that level, misgivings and market disturbances have at once begun" Noyes (1895, p. 575). The "legal minimum" refers to the provision in the Act of July 12, 1882, suspending the Treasury's issuance of gold certificates in exchange for coin deposit when the Treasury's reserves to redeem US notes fell below $100 million. This provision was repeated and extended in the Gold Standard Act of 1900. The redemption obligation covered both US notes and Treasury notes, and the Treasury was to have a reserve fund of $150 million for their redemption. When the fund fell below $100 million, the Treasury was to borrow to restore the Fund to the $150-million level. See Huntington and Mawhinney (1910, pp. 586–587, 610–614).
3. See Noyes (1895, pp. 588, 592; 1909, pp. 162, 232–233), Sprague (1910, pp. 141–142, 158, 179), Myers (1931, p. 378), Fels (1959, pp. 167, 185–186, 191–195), Simon (1960, pp. 32–33; 1968, p. 386), Friedman and Schwartz (1963, pp. 104–113), and Garber and Grilli (1986, p. 649). Also, using target-zone theory, Hallwood, MacDonald, and Marsh (1995) find that there were "large [dollar] devaluation expectations" during the 1890–1897 period.
4. In the 1890s, however, the bankers were no public-spirited agents, but rather were motivated by self-interest. Their provision of gold to the Treasury in 1895 involved the purchase of government bonds at 104½ (4.5% above face-value) at a time when comparable bonds were being transacted at 113½, and the bonds were marketed by the syndicate at 112¼ eleven days after the purchase. In four days of bargaining (in writing, by

dispatch), the syndicate maintained a tough bargaining stance and won its terms. Either the bankers were prepared to see the gold standard fall, or they were confident that the government would not let it fall (that is, would accept the terms of the syndicate). As Noyes (1895, p. 592) notes: "It was, however, perfectly plain that the administration had no choice but to accept the syndicate's proposition or suspend government specie payments."

5. The Act of August 27, 1894 imposed a tax of 2% on the "net profits or income" of corporations effective January 1, 1895, but in May 1895 the Supreme Court found the law to be unconstitutional. See Seidman (1938, pp. 1016–1020) and Stanley (1993, pp. 136, 299, n. 2).
6. Dutton (1984, p. 192) concludes from an econometric analysis that the Bank did not follow the "rules of the game" in its monetary policy. "Whether passive or active in the process, the Bank apparently acted as a buffer between reserve movements and money-supply changes. The rules would demand that it be an amplifier." However, the study deals only with policies other than gold-point manipulation.
7. As Dutton (1984, p. 178) notes, "They [the Bank's 'gold devices'] were usually used to retain or attract gold without resorting to extreme Bank-rate changes that would otherwise be necessary."
8. See Friedman and Schwartz (1963, pp. 154–156) and Goodhart (1969, pp. 111–114).
9. In a similar but milder vein, Dutton (1984, p. 178) states: "Their use was in a sense a violation of the rules, since the devices interfered with the free convertibility of gold at clearly specified rates of exchange."
10. For histories of these syndicates, see Noyes (1895, pp. 591–602; 1909, pp. 234–249), Simon (1960, 1968), and Friedman and Schwartz (1963, pp. 111–113).
11. See Brown (1914, pp. 113–114) and Cross (1923, p. 395).
12. Mathematically, let $u(R^*)$ denote the loss function relating u, the disutility from regime inefficiency, to R^*, the deviation of the exchange rate from the midpoint of the gold-point spread. Logical properties of the u function are (1) $u(0) = 0$; (2) $u(R^*) > 0$, for $R^* \neq 0$; (3) assuming u is continuously differentiable, $u' \gtreqless 0$ for $R^* \gtreqless 0$; (4) $u(R^*) = u(-R^*)$. Two fundamental u functions with these properties are the identity ($u = \left|R^*\right|$) and square ($u = \left|R^*\right|^2$). These loss functions involve disutility increasing at the rate of 1 and $2|R^*|$, respectively, as $\left|R^*\right|$ increases.
13. Suppose a given sample period of N observations. Then the experienced average (mean) loss from regime inefficiency is $\text{mean}(u) = (\Sigma u)/N$, which specifically is $\text{mean}\left|R^*\right|$ and $\text{mean}(\left|R^*\right|^2)$ for the two u functions adopted.
14. Conjoining the assumptions regarding $\mathrm{f}(R^*)$ and $u(R^*)$, the hypothetical average (expected) loss from regime inefficiency given the gold-point

parameter CX* is:

$$\mathrm{E}(u|\mathrm{CX}^*) = (1/\mathrm{CX}^*) \cdot \int_0^{\mathrm{CX}^*} U(R^*)\mathrm{d}R^* = [U(\mathrm{CX}^*) - U(0)]/\mathrm{CX}^*$$

where U is a particular integral of *u*. Intuitively, the first equality defines the expected value of *u* as the sum (integral) of all values of *u* in the interval [0, CX*], that is, all positive values of *u*, divided by the width of the interval (CX* – 0). (Because the function *u* is symmetrical, only the interval [0, CX*] need be considered.) The second equality provides the formula to calculate E(*u*|CX*).

For the identity and square functions, the respective U functions are $(R^*)^2/2$ and $(R^*)^3/3$, and computed values of E(*u*|CX*) ("integration formulas") are CX*/2 and $(\mathrm{CX}^*)^2/3$. Intuitively, the average value of $u(R^*)$ over [0, CX*] may be approximated by *u*(CX*/2), which correctly yields CX*/2 for the identity loss function. For the square function, however, this approximation is an understatement, $(\mathrm{CX}^*)^2/4$ rather than $(\mathrm{CX}^*)^2/3$, because the function increases at an increasing rate.

15. Because the denominator of the ratio is positively correlated with the width of the gold-point spread, the bias inherent in mean$|R^*|$ (and in the mean of any monotonic function of $|R^*|$) as a measure of efficiency—arising from neglect of the influence of the spread in constraining $|R^*|$—is corrected.
16. For the identity function, the efficiency ratio is:

$$100 \cdot \left[\mathrm{mean}|R^*|/(\mathrm{CX}^*/2)\right] = 200 \cdot \left[\mathrm{mean}|R^*|/(\mathrm{CX}^*)\right]$$

For the square function, the ratio is:

$$100 \cdot \left[\mathrm{mean}\left(|R^*|^2\right)/\left\{(\mathrm{CX}^*)^2/3\right\}\right] = 300 \cdot \left[\mathrm{mean}\left(|R^*|^2\right)/(\mathrm{CX}^*)^2\right]$$

17. In Officer (1993, p. 118) this efficiency ratio for 1925–1931 is mistakenly halved—my error. The analysis that follows there is too strong.

References

Brown, Harry Gunnison. *International Trade and Exchange*. New York: The Macmillan Company, 1914.

Clark, Truman A. "Violations of the Gold Points, 1890–1908." *Journal of Political Economy* 92 (October 1984), pp. 791–823.

Cole, Arthur H. "Seasonal Variation in Sterling Exchange." *Journal of Economic and Business History* 2 (November 1929), pp. 203–218.

Cross, Ira B. *Domestic and Foreign Exchange*. New York: The Macmillan Company, 1923.

Dutton, John. "The Bank of England and the Rules of the Game under the International Gold Standard: New Evidence." In Michael D. Bordo and Anna J. Schwartz, eds. *A Retrospective on the Classical Gold Standard, 1821–1931*. Chicago: University of Chicago Press, 1984, pp. 173–195.

Einzig, Paul. *The History of Foreign Exchange*, 2nd edn. London: Macmillan and Co., 1970.

Fels, Rendigs. *American Business Cycles, 1865–1897*. Chapel Hill: University of North Carolina Press, 1959.

Friedman, Milton and Anna Jacobson Schwartz. *A Monetary History of the United States 1867–1960*. Princeton, NJ: National Bureau of Economic Research, Princeton University Press, 1963.

Garber, Peter M. and Vittoriò U. Grilli. "The Belmont-Morgan Syndicate as an Optimal Investment Banking Contract." *European Economic Review* 30 (June 1986), pp. 649–677.

Goodhart, Charles A. E. *The New York Money Market and the Finance of Trade, 1900–1913*. Cambridge, MA: Harvard University Press, 1969.

Hallwood C. Paul, Ronald MacDonald, and Ian W. Marsh. "Credibility and Fundamentals: Was the Gold Standard a Well-Behaved Target Zone?" In Tamim Bayoumi, Barry Eichengreen, and Mark P. Taylor, eds. *Modern Perspectives on the Gold Standard*. Cambridge: Cambridge University Press, 1995.

Huntington, Andrew T. and Mawhinney, Robert J., eds. *Laws of the United States Concerning Money, Banking, and Loans, 1778–1909*. Senate Document No. 580, 61st Cong., 2nd session. Washington, DC: Government Printing Office, 1910.

Moggridge, Donald E. *British Monetary Policy 1924–1931: The Norman Conquest of $4.86*. Cambridge: Cambridge University Press, 1972.

Morgenstern, Oskar. *International Financial Transactions and Business Cycles*. Princeton: National Bureau of Economic Research, Princeton University Press, 1959.

Myers, Margaret G. *The New York Money Market, vol. I, Origins and Development*. New York: Columbia University Press, 1931.

Noyes, Alexander D. "The Treasury Reserve and the Bond Syndicate." *Political Science Quarterly* 10 (December 1895), pp. 573–602.

______. *Forty Years of American Finance*, 2nd edn. New York: G. P. Putnam's Sons, 1909.

Officer, Lawrence H. "Gold-Point Arbitrage and Uncovered Interest Arbitrage Under the 1925–1931 Dollar-Sterling Gold Standard." *Explorations in Economic History* 30 (January 1993), pp. 98–127.

______. *Between the Dollar-Sterling Gold Points: Exchange Rates, Parity, and Market Behavior*. Cambridge: Cambridge University Press, 1996.

Sayers, Richard S. *Bank of England Operations, 1890–1914*. London: P. S. King & Son, 1936.

Seidman, Jacob S. *Seidman's Legislative History of Federal Income Tax Laws, 1938–1861.* New York: Prentice-Hall, 1938.

Simon, Matthew. "The Hot Money Movement and the Private Exchange Pool Proposal of 1896." *Journal of Economic History* 20 (March I960), pp. 31–50.

______. "The Morgan-Belmont Syndicate of 1895 and Intervention in the Foreign-Exchange Market." *Business History Review* 42 (Winter 1968), pp. 385–417.

Sprague, Oliver M. W. *History of Crises Under the National Banking System*. Senate Document No. 538, 61st Cong., 2nd session. Washington, DC: Government Printing Office, 1910.

Stanley, Robert. *Dimensions of Law in the Service of Order: Origins of the Federal income Tax, 1861–1913.* New York: Oxford University Press, 1993.

Yeager, Leland B. *International Monetary Relations: Theory, History, and Policy*. New York: Harper & Row, 1976.

CHAPTER 22

U.S. Specie Standard

Republished with permission of Elsevier, from *Explorations in Economic History*, Vol. 39, No. 2 (April 2002), pp. 113–153; permission conveyed through Copyright Clearance Center, Inc.

22.1 The U.S. Specie Standard, 1792-1932: Some Monetarist Arithmetic

Critical to research on the monetary history of the United States is availability of a monetary-base series that is consistent, complete in coverage, and continuous over a long period. It is also important to have a balance-of-payments series with these same properties. Furthermore, the balance-of-payments series should be "monetary" in nature, reflecting the intimate relationship between the monetary base and balance of payments. Notwithstanding the pioneering research of Milton Friedman and Anna J. Schwartz, and the follow-up work of their students and others, these series do not exist. The main objective of this article is to develop these monetary-base and balance-of-payments series. The series can be used for new historical explorations and also for possible amendments of hitherto unchallenged results of previous investigations. Some examples are provided in the article, and the series are tabulated for further use by researchers.

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_22

When the First Bank of the United States opened for business on December 12, 1791, the United States was effectively on a specie standard, based predominantly on the Spanish silver dollar. The Mint Act of 1786 established a bimetallic standard with domestically produced coin, but this act had not been put into effect. The specie standard was formalized into legal bimetallism (Mint Act of April 2, 1792) and then gold monometallism (Act of June 22, 1874), and it remained the norm for the country until March 6, 1933, when President Roosevelt prohibited banks from paying out gold. Successive congressional and presidential action over the next 10 months eliminated both the specie standard and any mechanism for a return to it. By contrast, during the period 1792–1932, deviations from a specie standard and fixed exchange rate—that is, paper standards and floating exchange rates—were temporary aberrations.

This 141-year period witnessed three episodes of central banking, two Independent Treasury Systems, the classic pre-1914 gold standard, and occasional suspensions of specie payments. The comparative macroeconomic performance of logically determined subperiods composing 1792–1932 is the subject of this study. A generalized exchange market pressure model is used, and annual data series are developed to fit the model, also to examine monetary pyramiding and price and income behavior. Foremost among these series is the monetary base.

The famed Friedman and Schwartz (hereafter, FS) (1963, 1970) series of the monetary base for 1867–1932 is adjusted in light of a somewhat different methodology and is extended back to 1789. Consideration is also given to the work of Rutner (1974), who provides a monetary-base series in the FS tradition for 1833–1860, and Temin (1969), who generates a series autonomously for 1820–1857. Then the monetary balance of payments, consistent with the new monetary base, is generated for the full 1790–1932 period. The monetary-base and balance-of-payments series are presented as fundamental data contributions, beyond the analysis to which they are put in this study.

The methodology of the historical monetary base is discussed in Sect. 22.1.1. Whether or not the First and Second Banks of the United States were central banks seriously affects both the base and payments series, and this issue is considered in Sect. 22.1.2, leading to separation of 1792–1932 into subperiods (Sect. 22.1.3). The new monetary-base series is generated in Sect. 22.1.4 and presented in Sect. 22.1.5. Comparisons with the FS, Rutner, and Temin series, including amendments to historical findings, follow in Sect. 22.1.6. The monetary balance of payments is generated in Sect. 22.1.7. The new monetary-base and balance-of-payments data, along with specially developed series of price, income, and other variables, are put to use in a comparative evaluation of the performance of central-banking and other periods spanning 1792–1932 (Sect. 22.1.8). Following conclusions (Sect. 22.1.9), an appendix provides details on data sources and construction of variables; the text is devoted purely to analysis.

22.1.1 Methodology of the Historical Monetary Base

The importance of the monetary base is twofold. First, the money supply is the product of the money multiplier and the monetary base, with the multiplier being an explicit function of the commercial-banks' reserve/deposit ratio and the nonbank-public's currency/deposit ratio. This formulation is one of the great accomplishments of FS (1963, pp. 776–798), and they, followed by many imitators, use it repeatedly in their history to delineate the absolute and relative importance of the three determinants in changes in the money supply. Second, the monetary base is closely related to the monetary balance of payments, with a payments imbalance constituting the effect of international transactions on the monetary base. More generally, the monetary base and balance of payments, together with the exchange rate, combine to define exchange market pressure in the foreign-exchange market.

The monetary base is composed of all assets that are actual or potential reserves for the consolidated commercial-banking system. To make the definition operational, six questions must be answered:

1. Who holds the monetary base?
2. What are the assets that constitute the base?
3. For each asset separately, what is the time period for which it is included in the base?
4. What should be the dating pattern of the monetary-base series?
5. In what money should the base be denominated?
6. What data should be used, what interpolative techniques for missing data points, and under what circumstances is information so poor that, for example, a legitimate asset should be omitted from, or an illegitimate holder should be included in, the base on statistical grounds?

Monetary-base developers (and users) can reasonably differ on answers to each of these questions, depending on their objectives and the criteria that they use. The current study differs from the work of predecessors in making fully explicit these objectives and criteria.

The objectives are (a) to achieve consistency over a long duration, 1792–1932; (b) to consider the United States as on a virtual, if not actual, specie standard throughout the time span; and (c) to ensure compatibility with the monetary balance-of-payments measure. The criteria are (i) to apply strictly the definition of the monetary base and operate in accord with the objectives in answering questions 1–5, and (ii) to use all available information to maximum effectiveness in answering question 6. Administering criterion (ii) inevitably involves considerable judgment, and again reasonable researchers can differ in their decisions. The advantages of the current study over predecessors in this respect emanate from the work of FS and their students, the existence of specialized studies pertinent to the monetary base written since their time, and spreadsheet/statistical programs that were not available to FS.

22.1.2 Were the First and Second Banks Central Banks?

FS (1970) do not address the issue of whether the First and Second Banks were central banks. However, in showing data for these banks separate from state banks, they leave the question open. For the current study, the pertinent central-banking criterion is whether the Banks' note circulation (and, by extension, non-Treasury deposit liabilities) served as actual or potential reserves for the state banking system and hence constituted part of the monetary base.[1] While the question has not been directly addressed for the First Bank, many have answered in the affirmative for the Second Bank.[2] Yet it would be a reasonable position that, given the controversial nature of these institutions and the long tradition of considering specie as ultimate money, the Banks' liabilities were considered just ordinary money. Fortunately, a variety of empirical evidence exists on the matter.

First, Fenstermaker (1965, p. 43) and Rutner (1974, p. 25, n. 1) note, for the First and Second Banks, respectively, that Bank notes were sometimes included with specie in the statements of state banks.[3] Second, Fenstermaker (1965, pp. 11–12, 69–76) synopsizes the entire history of the Second Bank in terms of its credit contraction/expansion with multiple effect on credit contraction/expansion of the state banking system.[4] Third, Engerman (1970, p. 726) and Rutner (1974, pp. 23–30, 121–146) show that the nonbank public considered Second Bank notes and deposits as substitutes for specie, the primary base money; and Rutner provides even stronger evidence for this treatment on the part of the state banks. Furthermore, the base-money characteristic of Bank note and deposit liabilities continued many months after February 1836 (the date of replacement of the Bank's federal charter with a Pennsylvania charter) and even after the Bank's initial suspension of specie payments (May 1837 to August 1838)—by Rutner's evidence, until "sometime in 1839," probably with the Bank's second suspension in October.[5]

The same reasons underlying the monetary-base property of Second Bank note and deposit liabilities apply to those of the First Bank, and hence the positive empirical findings for the Second Bank may be extrapolated to the First Bank. Each Bank was a balance-sheet giant in comparison to contemporary state banks, and, as national institutions, each had branches in the major commercial cities of the country.[6] Each was the fiscal agent of the government and served as a major (First Bank) or sole (Second Bank—to 1833) despository of the Treasury. These circumstances generated a large and steady stream of state bank notes (and checks) to the Banks, which generally presented them regularly to the state banks for redemption in specie. These banks, in turn, could avoid specie loss by presenting the Bank with the Bank's notes and drawing down its deposits at the Bank. Therefore, Bank note and deposit assets were considered by the state banks as part of reserves.

The Banks' redemption practice was a technique of monetary control that was fostered by the conservative credit policy of the First Bank and by the conscious regulation of the state banks on the part of the Second Bank under President Nicholas Biddle. When the First or Second Bank chose not to redeem its state bank notes, it became a still greater creditor of these banks, thereby enhancing future control. Hammond (1947) argues that this regulatory power—different from modern central banking in the creditor rather than debtor status of the central bank with respect to commercial banks—was "simpler, more direct, and perhaps more effective than those of the Federal Reserve Banks" (p. 2).[7]

The notes of the Banks were clearly superior to state bank notes. By federal charter, Bank notes were legal tender for all payments to the government. Combined with interstate banking, this gave rise to universal acceptability in the private sector—not a characteristic of state banks at the time. The conservative note-issuance policy of the First Bank and the effectiveness of the Second Bank in reducing the deviation of domestic exchange rates from parity were additional elements in producing Bank note issue that compared favorably to the specie stock in uniformity and cost of transfer.[8]

22.1.3 Delineation of Subperiods

As suggested in the introduction, and consistent with the "contingent-rule gold-standard" concept developed by Bordo and Kydland (1995) and Bordo and Rockoff (1996), there is a real sense in which the United States was on a metallic standard throughout 1792–1932, with deviations from paper-currency convertibility deemed to be, and in fact, temporary. Nevertheless, subperiods of interest may be distinguished, primarily by identification of a monetary authority (First and Second Banks, Federal Reserve Banks, Independent Treasury) and secondarily by the longest suspension of specie payments (greenback period) and the "classic" gold standard that followed.[9]

With the First Bank in operation from December 12, 1791, to the expiration of its charter on March 4, 1811, 1792–1810 is naturally the first period of central banking. The interregnum between the First and Second Banks is 1811–1816, a period of issuance of the first Treasury currency component of the monetary base (Treasury notes) and, beginning August 30, 1814, the first major suspension of specie payments. The Second Bank opened for business on January 7, 1817, and was treated as a central bank by the state banking system into 1839, yielding 1817–1838 as the second period of central banking.[10] Another interregnum, 1839–1846, includes paper standards (parts of 1839–1842 over much of the country) and the aborted first Independent Treasury System (July 4, 1840, to August 13, 1841).

The years 1847–1861 constitute the (second) Independent Treasury System, which began on January 1, 1847, when all payments to the Treasury were by law in specie or Treasury notes (not state bank notes). From April 1, 1847, payments from the Treasury were similarly made. Throughout this period, funds were kept within the government; banks were not used as depositories. The Act of August 5, 1861, began erosion of the policy, permitting proceeds of the first substantial Civil War loan to be deposited in state banks.

On December 30, 1861, virtually all banks ceased converting their notes and deposits into gold coin, and the Treasury suspended the right of holders of its demand notes to redeem them in gold. Resumption occurred on January 1, 1879, defining 1862–1878 as the greenback period. After the classic gold standard, 1879–1913, the third period of central banking began with the creation of the Federal Reserve System by the Act of December 23, 1913. The United States abandoned the gold standard on March 6, 1933, making 1932 the specie-standard's last full year of operation.

22.1.4 The New Monetary Base, 1789–1932: Construction

22.1.4.1 Structure

In the FS tradition, the monetary base consists of all assets—gold or specie, nongold metallic money, (paper) currency, and deposits—that the consolidated private banking system can use as reserves either actually (these assets held by banks) or potentially (these assets held by the public). By definition, assets in (domestic) circulation are the sum of assets held by the banks and by the public. The monetary base is provided by "outside" agents, and increases or decreases in components of the base occur via transactions of the "inside" entities (the banks and public) with the outside. The outside agents are (1) the foreign sector (affecting the specie stock via international transactions), (2) the nonmonetary sector (altering the specie stock via production of bullion and consumption of bullion or coin), (3) the Treasury (producing nongold metallic money and paper currency but reducing the base by using specie as backing for issued currency), and (4) the central bank (providing paper currency and deposits, using specie as reserves for same).[11] Also incorporated are gold certificates (circulating warehouse receipts for gold deposits at the Treasury), lost currency, foreign-held currency, and nonunitary specie price of currency.

As the supply of base money (BASE_S), the monetary base is the sum of the net contributions of specie, the Treasury, and the central bank. The contribution of specie is the amount of specie in the country (commonly called the "specie stock") *minus* lost gold certificates. The gross contribution of the Treasury is its currency (excluding gold certificates) in official circulation *minus* lost currency *plus* nongold coin in circulation.[12] For the Treasury net contribution, there are two deductions: Treasury net specie (Treasury gross specie *less* Treasury gold held against gold certificates) and Treasury currency held by foreigners. Treasury gold held against gold certificates equals these certificates in official circulation: the sum of certificates in circulation and certificates lost. The gross contribution of the central bank is its currency in official circulation *minus* lost currency *plus* non-Treasury domestic deposit liabilities. The central-bank net contribution is obtained by subtracting its specie and its currency held by foreigners.

Monetary variables are expressed in millions of "gold dollars" (incorporating "specie dollars" prior to 1860), except that the components of the gross contribution of the Treasury and the central bank are in millions of paper dollars. To convert to gold dollars, the gross contributions are multiplied by the specie price of currency (par of unity).[13]

22.1.4.2 Comparison with Other Historical Monetary-Base Series

Composition of base. The new monetary base centers on the *net* liabilities (fiduciary contributions to the base) of the authorities, which measures the Treasury and central-bank contributions given the specie stock. There are no precedents for this partitioning of the historical monetary base. The usual breakdown of the historical base focuses on the *gross* liabilities of the *combined* authorities; the specie stock is replaced by specie in circulation (specie stock *less* Treasury and central-bank specie), while Treasury and central-bank currency are combined. This composition—found in FS (1963, pp. 130, 179, 704–722, 735–744) and Rutner (1974, pp. 151–183) as well as in Kindahl (1961, p. 40)—minimizes the role of specie and does not delineate the contributions of the respective authorities to the base. However, the monetary-base *aggregate* is not affected by these alternative partitions.

Classification of gold certificates. Circulation of gold certificates (first issued in nontrivial amount in 1866) is subsumed in the gold stock and therefore in the contribution of that stock to the monetary base. This placement is in accord with the net-liabilities format and enhances the role of specie relative to the Treasury. It is in contrast to the FS treatment of gold certificates as currency. However, FS (1963, p. 25, n. 12) themselves provide two justifications for the former procedure: the pure warehouse-receipt nature of the certificates and (during the greenback period) the market's refusal to recognize a premium on the certificates below that for gold itself. Again, the monetary-base aggregate is invariant to where gold certificates are placed.

Dating pattern of series. Uniform end-of-year dating is adopted, for consistency over the 1789–1932 time span and for compatibility with the monetary balance-of-payments series.[14] FS provide end-of-year figures only from 1907, while Rutner has 6 years that lack this dating, but their objective is rather to maximize the frequency of observations subject to a given level of data reliability. Temin's series pertains to the end of the fiscal (rather than calendar) year, because that is the timing of the flow data underlying his series.

Definition of the public. Temin includes both the Treasury and the Second Bank in the public. The result is that the monetary base reduces to the specie stock. Because the Treasury did create money during the antebellum period (recognized but not emphasized by Temin), which money was used as bank reserves, Rutner is justified in treating the Treasury as an outside agent. Also, Rutner's decision to classify the Second Bank as a central bank was supported in Sect. 22.1.2. Therefore, it is reasonable to follow Rutner in rejecting Temin's additions to the public.

The FS monetary base *includes* not only Treasury and Federal Reserve currency held by the domestic public and banks but also such currency held by the foreign public and banks (FS 1963, p. 778; 1970, pp. 58–60). However, the FS base *excludes* U.S.-issued gold and silver coin held by foreigners. While Garber (1986) is correct in observing this inconsistency in the definition of the public, FS are simply following official data on currency and coin in circulation. It is the reporting of currency data by issuers rather than holders of money that leads to the inconsistency in the FS base. Indeed, FS note that "in principle" and "ideally defined," foreign-held dollars should be excluded from the base. The FS (and Garber) ideal is followed in the current study, because data do exist to exclude foreign-held dollars from the base.

Denomination of base. FS (1963) sum gold-dollar-denominated and paper-dollar-denominated components of the monetary base during the greenback period. They are well aware that this arithmetic is analogous to adding apples and oranges: "Treating one greenback dollar as equal to one gold dollar... [is], strictly speaking, meaningless: it is like adding current Canadian or Hong Kong dollars to U.S. dollars on a one-to-one basis" (FS 1963, pp. 27–28). The same issue arises during May 1837 to August 1838, when the Second Bank suspended specie payments and its notes depreciated in terms of gold. The depreciation of Second Bank money is ignored by Rutner; but FS justify their simple summation of gold and depreciated dollars on two grounds: "[It] is done…in every other summary of monetary statistics for the greenback period we know of" (FS 1963, p. 28), and the necessary correction declines over time with the decrease in the gold premium.

Because the United States is considered in essence to be on a specie standard throughout 1789–1932, and because consistency over time is desired, the new monetary base is uniformly expressed in gold (or specie) dollars. This is done via deflation of base components that traded at a discount in terms of gold during periods of paper-currency depreciation.[15] The specie price of currency for the central bank is non-unity only for 1837, and that for the Treasury is non-unity only for the greenback period. There was no central bank during the greenback period, and the depreciation of Second Bank liabilities during 1837–1838 did not affect the par value of Treasury currency.

Attention to lost currency. Official currency in circulation, used in the FS base, includes "currency irretrievably lost, destroyed, in collections, or otherwise so disposed as never to be presented for redemption" (Laurent 1974, p. 213, n. 1); such "lost currency" is deducted in constructing the new monetary base. FS (1963, pp. 442–443, n. 20) are aware of the issue and estimate the loss for national bank notes at about 0.1% per year, but they do not adjust their monetary base for lost currency. It may be that they judged the correction to be quantitatively unimportant based on their finding for national bank notes, or perhaps they did not see how to estimate the deduction for other forms of currency.

Treatment of state bank notes. State bank notes, included in the FS base to mid-1878 (see FS 1963, pp. 722, 724, 808), are clearly not high-powered money, are removed from the FS base by Joines (1985, p. 348), and are not a component of the Rutner base. They are excluded from the new monetary base. FS neglect to make this correction as well, probably because they deemed it to be of minor quantitative importance.

Treatment of national bank notes. FS (1963, pp. 20–23, 50, 780–782) include national bank notes in the monetary base—reasonably because this currency served as a reserve for state banks and was legal tender for Treasury transactions (with exceptions). However, the current study places national bank notes in Treasury currency (and therefore in the monetary base) only from 1874, for reasons stated by FS themselves. Legally, a reserve requirement had been imposed on both notes and deposits of national banks, beginning with the first National Banking Act (February 25, 1863). Only with the Act of June 20, 1874, was the reserve requirement removed from national bank notes, while being retained on deposits. Furthermore, this act—and not the, earlier, National Banking Acts—provided for Treasury redemption of national bank notes in U.S. notes at par (based on a fund to which banks contributed 5% of their note issue, countable toward their reserves on deposits). Empirically, there was the potential, and in at least one instance (early 1873) the actuality, of national bank notes trading at a discount for U.S. notes.

Whether or not national bank notes should be included in Treasury currency and therefore in the monetary base prior to 1874, as done by FS, is a matter of judgment. On the side of inclusion is the fact that national bank notes were backed more than fully (111% of value of notes issued) by government bonds deposited with the Treasury and therefore can be construed as an indirect obligation of the government, that is, as Treasury currency at one remove. On the side of exclusion, viewed in this study as preponderant, are the existence of a reserve requirement, the absence of a redemption fund, and the trading of national bank notes at a discount in terms of greenbacks. As stated by FS, "[In] the period before 1874…[national bank] notes were more nearly identical with deposits than with the notes issued by the Treasury," and "[To] treat national bank notes as part of the currency obligations of the monetary authorities…is of questionable appropriateness for the first few years covered by our series" (FS 1963, pp. 781–782).[16]

Selection of data. In respect of data used, the new monetary base is closer to FS than to Rutner or Temin. The antebellum specie stock is constructed via a new technique and with substantially different data from those of Rutner and Temin. Also, Treasury gold and Treasury notes during the antebellum period have different data sources from those of Rutner. By contrast, the FS gold stock, specie stock, gold certificates, and nongold coin series are accepted and extended back to 1860. Prior to 1874, only part of Treasury currency is consistent with FS. From 1874, the entirety of Treasury currency (and of Federal Reserve liabilities, from 1914) has data compatible with FS.

The result is that the new monetary base is different from the FS, Rutner, and Temin series. Components of the net-liabilities composition of the new base are discussed below. Subsequently, Sect. 22.1.5 presents empirically the contributions to the base emanating from both the net-liabilities and an alternative breakdown, authorities' net assets. Then Sect. 22.1.6 shows just how different the new base is from its predecessors.

22.1.4.3 Components of Net-Liabilities Composition of the Monetary Base

Specie stock and nongold coin. Prior to 1860, data on the specie stock include both gold and silver, although by the late 1850s silver is in the form only of domestic subsidiary coinage.[17] From June 30, 1860, official specie-stock series are limited to gold, consisting of domestic gold coin in circulation and gold in all forms (domestic coin, foreign coin, and bullion) in the Treasury or Federal Reserve. From that date, nongold coin (standard silver dollars, subsidiary silver coin, and minor coin) became separate official series. The specie stock for the new monetary base follows the official line—gold and silver to the end of 1859 and gold alone thereafter. The specie stock, its distribution, and nongold coin circulation agree with the corresponding FS series.

Contribution of Central Bank to monetary base. Both the First and Second Banks issued not only banknotes (payable in specie on demand) but also postnotes (payable in specie on demand at a specified future date after issuance). For the First Bank, postnotes are included in central-bank currency, because (1) they were issued regularly only by the main office and in the ordinary course of business, and (2) "[Total] note circulation was deliberately restricted to guard specie" (Wettereau 1937, p. 283); there was never a question of suspending specie payments. For the Second Bank, postnotes are excluded. The Second Bank first issued postnotes in March 1837 decidedly not in the ordinary course of business, while specie payments were suspended, in an attempt to obtain specie.[18]

Contribution of Treasury to monetary base. The FS composition of Treasury currency is followed in its inclusion of national bank notes (from 1874), silver certificates (receivable for all payments to the Treasury from inception in 1878, and a legal reserve for national banks by the Act of July 12, 1882), Treasury notes of 1890 (a full legal tender), U.S. notes (greenbacks, first issued in 1862, a legal tender with exception for certain payments to the Treasury), fractional currency (instituted in 1863, a substitute for subsidiary silver coin), and certain Civil War issues designated as "other U.S. currency" in official statistics: old demand notes (payable for all public dues, made legal tender by the Act of March 17,1862), Treasury notes of 1863, and compound-interest notes (both interest-bearing but legal tender on the same basis as U.S. notes).

However, Treasury currency differs from the FS concept in two respects. First, gold-certificate circulation is subsumed in the gold stock, in contrast to the FS treatment of gold certificates as currency. Second, 3% certificates, issued after the Civil War, were a legal reserve for national banks and so are included in Treasury currency.[19] Also, the FS concept must be broadened in two respects for extension prior to 1867. First, postage currency, issued for nearly a year beginning July 1862 and replaced by the fractional currency, is included in Treasury currency (in fact, the two types of currency are intermixed in official statistics).[20] Second, Treasury notes, issued between 1812 and 1861, also are included in Treasury currency; interest-bearing, they had the same legal-tender characteristic as did Bank of United States notes, were used as bank reserves, and (in small denominations) even served as hand-to-hand currency.[21]

22.1.5 The New Monetary Base, 1789–1932: Presentation

22.1.5.1 Net-Liabilities Breakdown

The new monetary base for 1789–1932 is listed in Table 22.1. The contributions of the specie stock, Treasury, and central bank are presented as period averages in Table 22.2. The contributions of the Treasury and central bank can be interpreted as the reduction in the monetary base should the Treasury or central bank be reclassified to the private sector. In particular, treating the First and Second Banks as commercial banks would reduce the monetary base by 18 and 20%, on average. In principle, the contribution of the Treasury or the central bank can be negative, and in fact that of the Treasury is negative during 1849–1857 and 1917–1932, averaging −\$813 million during 1914–1932. FS (1963, pp. 391–393, 399) consider a hypothetical policy of a Federal Reserve \$1 billion open-market purchase of securities in 1930 or 1931, which would have moderated, and possibly prevented, the crises that led to the Great Depression. All the while, the Treasury was immobilizing a *greater* amount of gold; its net contribution to the base was *negative* \$1167,

Table 22.1 Monetary Base and monetary balance of payments, 1789–1932 (millions of gold dollars)

Year	*Monetary Base (end of year)*	*Balance of Payments*[a]		*Year*	*Monetary Base (end of year)*	*Balance of Payments*[a]	
		Direct	*Indirect*			*Direct*	*Indirect*
1789	9			**1861**	317	−5	−8
1790	16		7	**1862**	416	−41	−47
1791	18		2	**1863**	482	−74	−84
1792	22		3	**1864**	421	−72	−82
1793	24		1	**1865**	506	−59	−69
1794	25		−3	**1866**	452	−45	−54
1795	21		−3	**1867**	399	−46	−53
1796	20		−1	**1868**	369	−46	−51
1797	18		−2	**1869**	410	−24	−28
1798	22		3	**1870**	442	−41	−48
1799	22		1	**1871**	436	−50	−58
1800	22		0	**1872**	419	−39	−47
1801	21		0	**1873**	426	−25	−32
1802	18		0	**1874**	700	−34	−42
1803	22		2	**1875**	672	−38	−43
1804	22		1	**1876**	700	−12	−17
1805	23		1	**1877**	732	2	−2
1806	23		2	**1878**	746	2	1
1807	24		0	**1879**	867	39	78
1808	23		0	**1880**	1001	87	67
1809	25		0	**1881**	1113	50	63
1810	26		0	**1882**	1148	4	−28
1811	21		1	**1883**	1180	−6	23
1812	25		1	**1884**	1210	−1	−10
1813	28		1	**1885**	1202	−3	11
1814	35		1	**1886**	1219	5	8
1815	48		6	**1887**	1285	28	37
1816	33		−1	**1888**	1315	−13	−27
1817	38		−1	**1889**	1333	−28	−42
1818	35		−1	**1890**	1420	−38	−7
1819	30		−1	**1891**	1483	−35	−43
1820	29	−1	1	**1892**	1502	−44	−63
1821	33	−3	2	**1893**	1598	−46	−16
1822	32	−6	0	**1894**	1498	−17	−83
1823	31	1	2	**1895**	1441	−54	−72
1824	34	−2	−2	**1896**	1501	−17	43
1825	41	1	3	**1897**	1569	75	1

(continued)

Table 22.1 (continued)

Year	*Monetary Base (end of year)*	*Balance of Payments*[a]		*Year*	*Monetary Base (end of year)*	*Balance of Payments*[a]	
		Direct	*Indirect*			*Direct*	*Indirect*
1826	38	2	1	**1898**	1732	77	140
1827	41	–1	–1	**1899**	1821	23	9
1828	44	2	2	**1900**	2025	3	28
1829	44	4	2	**1901**	2099	7	0
1830	48	5	2	**1902**	2195	–1	11
1831	56	–5	–5	**1903**	2309	6	19
1832	51	6	4	**1904**	2413	–12	–36
1833	55	6	0	**1905**	2505	8	7
1834	61	14	11	**1906**	2715	59	104
1835	76	5	4	**1907**	3021	68	97
1836	73	8	7	**1908**	3054	13	–25
1837	74	–3	–3	**1909**	3084	–63	–88
1838	90	9	7	**1910**	3161	–13	9
1839	76	13	13	**1911**	3238	20	24
1840	75	–1	–1	**1912**	3320	–10	23
1841	74	–4	–3	**1913**	3403	–35	–35
1842	82	6	6	**1914**	3386	–18	–168
1843	86	14	12	**1915**	3788	288	416
1844	81	–2	–2	**1916**	4413	516	461
1845	78	–2	–2	**1917**	5436	219	250
1846	90	11	9	**1918**	6302	–216	–219
1847	102	6	6	**1919**	6504	–284	–287
1848	97	–4	–4	**1920**	6670	–108	–125
1849	94	–1	–2	**1921**	5668	610	630
1850	118	–13	–19	**1922**	5804	155	174
1851	142	–31	–36	**1923**	6029	253	259
1852	165	–30	–36	**1924**	6340	248	270
1853	191	–29	–34	**1925**	6529	–39	–54
1854	202	–44	–45	**1926**	6481	61	67
1855	201	–47	–47	**1927**	6621	–57	–33
1856	205	–49	–50	**1928**	6599	–225	–208
1857	216	–45	–46	**1929**	6485	152	158
1858	244	–45	–45	**1930**	6678	364	367
1859	235	–57	–56	**1931**	7287	–239	–237
1860	250	–21	–43	**1932**	7673	160	148

[a]Net specie imports

Table 22.2 Contributions to monetary base: net-liabilities breakdown (period means—end of year)

Period	*Millions of Gold Dollars*			*Percentage of Monetary Base*		
	Specie	*Treasury*	*Central Bank*	*Specie*	*Treasury*	*Central Bank*
1791–1810	18	–	4	82	–	18
1811–1816	25	7	–	82	18	–
1817–1838	38	1	9	79	1	20
1839–1846	76	5	–	94	6	–
1847–1861	185	0	–	101	−1	–
1862–1878	149	364	–	31	69	–
1879–1913	931	983	–	47	53	–
1914–1932	3708	−813	3141	61	−11	50

\$1091, and \$1359 million during 1929–1931. Treasury action to increase its monetary-base contribution to *zero* was a logical alternative to Federal Reserve policy.

As expected, the share of specie in the monetary base is highest during the Independent Treasury period and lowest during the greenback period. Perhaps surprising is that only during the latter period and 1879–1913 does the Treasury make a large relative contribution to the monetary base.

22.1.5.2 Assets Breakdown

The composition of the monetary base that leads directly to the balance of payments centers on the assets of the combined Treasury and central bank (with new variables measured in millions of gold dollars). The monetary base ($BASE_S$) is the sum of (1) the specie stock (SPST), (2) net foreign assets (*excluding specie*) of the Treasury and central bank (NFA), and (3) the residual contribution of the Treasury and central bank to the base (RCON):

$$BASE_S = SPST + NFA + RCON \tag{22.1}$$

NFA consists of Treasury and central-bank currency held by foreigners (liabilities of the authorities, therefore with negative sign) *plus* central-bank net foreign assets other than foreign-held currency.

The residual contribution of the authorities (RCON) has two positive, and three negative, components. The positive terms are (i) Treasury currency in official circulation *less* currency lost *plus* nongold coin in circulation (from 1860, previously in the specie stock) and (ii) central-bank currency in official circulation *less* currency lost *plus* non-Treasury domestic deposits at the central bank, with both (i) and (ii) multiplied by the pertinent specie price of currency. These two terms represent fiat currency of the Treasury, central-bank domestic credit, and physical assets of these authorities (silver stock of the Treasury [from 1860] and premises of the central bank).

The negative components of RCON are lost gold certificates, Treasury net and central-bank specie, and central-bank net foreign assets other than foreign-held currency. Unlike in the net-liabilities composition, lost gold certificates are deducted from RCON rather than from the specie stock. With the specie stock rather than specie in circulation a component of the monetary base, Treasury net and central-bank specie must be subtracted from RCON. Central-bank net foreign assets other than foreign-held currency are deducted for inclusion in NFA, but foreign-held Treasury and central-bank currency are negative components of NFA rather than of RCON. These groupings are preparatory for balance-of-payments derivation in Sect. 22.1.7. The asset breakdowns of the historical base provided by previous authors—FS (1963, pp. 210–212, 796–798), Cagan (1965, pp. 333–339), and Bordo (1975, p. 511)—do not separate NFA and therefore do not link to the balance of payments.[22]

Period averages for the assets composition of the monetary base are shown in Table 22.3. The relative unimportance of NFA is noteworthy. The pattern of the specie contribution versus the two other components is the same as for Table 22.2, with the relative contribution of the specie stock a maximum under the Independent Treasury System and a minimum during the greenback period. In absolute terms, the specie stock expands more than six-fold during 1879–1913 over the greenback period and again four-fold during 1914–1932 over 1879–1913. The residual contribution of the authorities is only 19 and 22% under the First and Second Banks, respectively. Certainly, the First Bank was always a conservative institution, and the Second Bank could be described as such for a good part of its federally chartered existence. Even during the Federal Reserve period, the authorities' residual contribution is outweighed by the specie stock.

Table 22.3 Contributions to monetary base: assets breakdown (period means—end of year)

Period	*Millions of Gold Dollars*			*Percentage of Monetary Base*		
	Specie	*Treasury and Central Bank*		*Specie*	*Treasury and Central Bank*	
		Net Foreign Assets	*Residual Contribution*		*Net Foreign Assets*	*Residual Contribution*
1791–1810	18	0	4	82	−1	19
1811–1816	25	–	7	82	–	18
1817–1838	38	−1	11	80	−2	22
1839–1846	76	–	5	94	–	6
1847–1861	185	–	0	101	–	−1
1862–1878	149	–	364	31	–	69
1879–1913	936	−5	983	47	0	53
1914–1932	3726	−355	2665	62	−6	44

22.1.6 The New Monetary Base Versus Predecessors

Comparison of the new monetary base with predecessor series is instructive. The Temin (1820–1857), Rutner (1833–1859), and FS (1867–1932) series are obtained on a uniform year-end basis, compatible with the new base, via linear interpolation between adjacent figures closest to year end.[23] Temin's monetary base is the specie stock. So both the new monetary base and the new specie stock are compared to it, with the three series graphed in Fig. 22.1. For 1820–1857, the new specie stock averages 24% *below* the Temin series. The new base averages 12% *above* the Temin series during the period of the Second Bank (1820–1838) but 35% *below* it thereafter. The Temin series is smoother than the new base, but after 1838 it diverges sharply upward.

The new monetary base is graphed against the Rutner series in Fig. 22.2. The new base averages 10% *below* the Rutner series during 1833–1850, 13% *above* it during 1851–1854 (the only years when the new base exceeds Rutner), and 17% *below* it during 1855–1859. The two series have a broadly similar pattern until 1850.

The new monetary base is uniformly less than the FS base, as shown in Figs. 22.3–22.5. It averages 46.7% below the FS base during 1867–1873 (principally due to the exclusion of national bank notes), 2.5% below

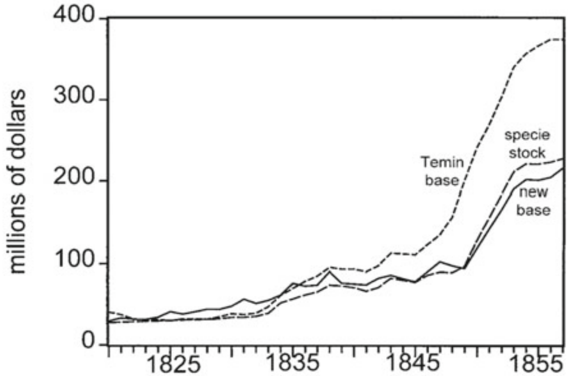

Fig. 22.1 New monetary base and specie stock versus Temin monetary base, 1820–1857

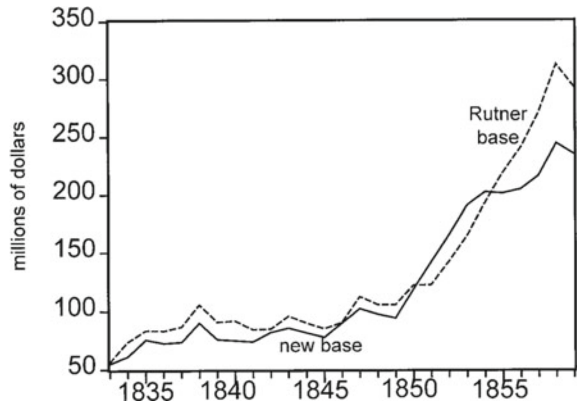

Fig. 22.2 New monetary base versus Rutner monetary base, 1833–1859

it during 1874–1897, 1.8% below it during 1898–1917 (when non-European foreign-held dollars are deducted), and 7.5% below it during 1918 1932 (when European-held dollars also are excluded). Only during 1867–1878 do the series diverge sharply. Afterward, they track each other very closely.

The new base suggests amendments to historical investigations of the determinants of the money stock. First, the new series is always below the

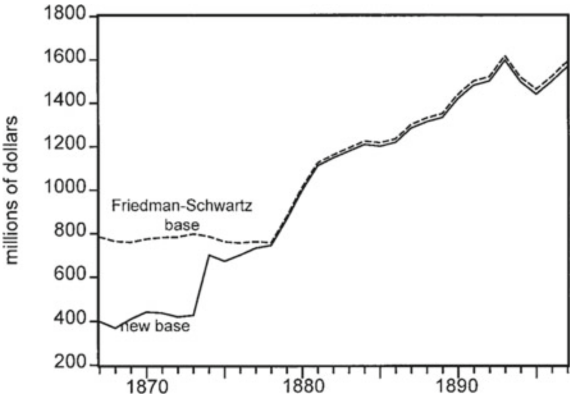

Fig. 22.3 New monetary base versus Friedman-Schwartz monetary base, 1867–1897

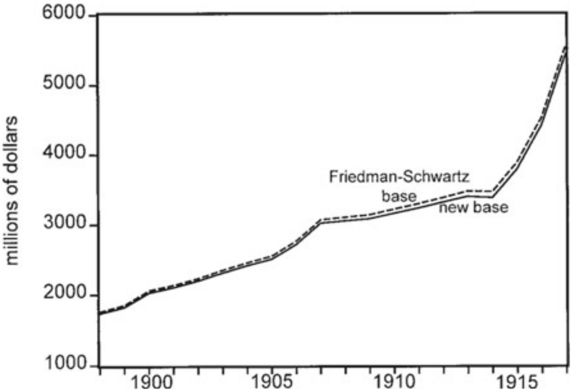

Fig. 22.4 New monetary base versus Friedman-Schwartz monetary base, 1898–1917

FS base. The implication is that, for a given period or point in time, the FS series would overestimate the role of the monetary base relative to the reserve/deposit and currency/deposit ratios, compared to results using the new base. In this respect, the *level* of the monetary base matters—pertinent because (except for 1867–1878) in percentage changes the new series is broadly similar to its predecessors. However, researchers typically

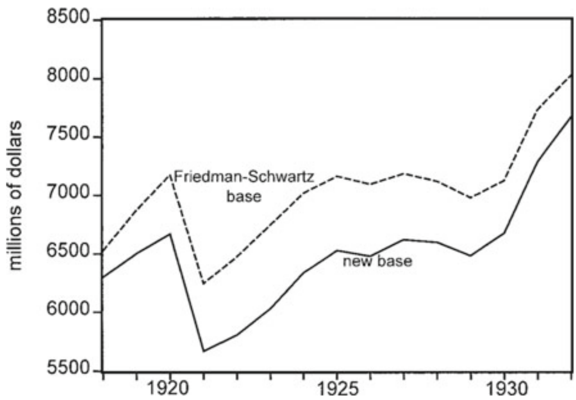

Fig. 22.5 New monetary base versus Friedman-Schwartz monetary base, 1918–1932

are concerned instead with *changes* in the money stock and in determinants of the money stock. Even here, the new base can make a difference. Some examples follow.

Temin (1969) observes, “The factor leading to an expansion of the monetary stock, then, was the rise in the stock of specie. The amount of specie in the country more than doubled in the quinquennium following 1832” (p. 77). Temin shows an official specie-stock series for this discussion, and the 1832–1837 increase is 184%, compared to a money-supply growth of 55%. His own specie series increases by 114% (116% on an end-of-year basis). By contrast, the new monetary base increases by only 44%—*less than the money supply expansion.*

FS (1963, p. 53) note “the mild and almost horizontal movement in high-powered money” from January 1867 to February 1879, with their monetary base changing at an annual average rate of −1.03% compared to 1.33% for the money supply. The new monetary base increases at an annual average rate of 4.18%, and *dominates* the other two determinants of the money supply rather than, as FS found, the reverse.[24] For July 1921 to August 1929, FS (1963, p. 275) describe the “change in high-powered money…[as of] minor importance for the period as a whole.” Their figures show annual average percentage increases of 4.6% for the money supply and only 1.3% for their monetary base. The latter figure compares to 0.7% for the new base (from average 1920–1921 to average 1928–1929). In this case, using the new base *strengthens* the FS finding.

22.1.7 The Monetary Balance of Payments, 1790–1932

The methodology for the monetary balance of payments was developed by Kemp: "Compute the net impact of [international] transactions on the U.S. money stock…Of all international transactions, the only ones that affect the money stock are those that affect some component of the monetary base"(Kemp 1976, p. 10; see also , 1975a, 1975b). In this light, the existing historical balance-of-payments series—North (1960, pp. 600, 605) for 1791–1860, Simon (1960, pp. 699–705) for 1861–1900, and Bureau of the Census (hereafter "Census" 1975, pp. 867–868) for 1874–1932—fall far short.

First, these series ignore net foreign assets of the Treasury and central bank, confining attention to specie transactions alone. Second, there are specific conceptual errors. The Bureau of the Census defines the balance of payments as the entire change in the gold stock, intermixing net production of gold—a purely domestic transaction—with net imports. Simon includes silver as well as gold in specie, thereby creating an inconsistency not only with official monetary data but also with the FS and new monetary bases. Third, North and Simon use official data, that measure net specie imports directly, whereas it is indicated below that an indirect computation provides the more-reliable series. Fourth, North uses inconsistent data from various sources to estimate the series for 1790–1819, and his interpolation method is opaque.

Let IMP denote annual *net* specie imports and PROD annual *net* specie production, the difference between gross production and nonmonetary consumption (import of coin or bullion, production of bullion, or melting of coin that is retained as bullion or used in arts and industry *minus* nonmonetary metal melted down and recoined). Consider the equation.

$$\Delta \text{SPST} = \text{IMP} + \text{PROD} \tag{22.2}$$

With SPST, and therefore ΔSPST, known and of a higher order of accuracy than IMP and PROD (see Appendix), Eq. (22.2) will necessarily hold only if either IMP or PROD is estimated residually. Suppose rather that *both* variables are constructed directly, with notation IMP^{dir} and PROD^{dir}, and let the residual $\text{RES} = \Delta\text{SPST} - (\text{IMP}^{\text{dir}} + \text{PROD}^{\text{dir}})$.

Then the indirectly estimated variables are (a) $\mathrm{IMP}^{\mathrm{ind}} = \Delta\mathrm{SPST} - \mathrm{PROD}^{\mathrm{dir}} = \mathrm{IMP}^{\mathrm{dir}} + \mathrm{RES}$ and (b) $\mathrm{PROD}^{\mathrm{ind}} = \Delta\mathrm{SPST} - \mathrm{IMP}^{\mathrm{dir}} = \mathrm{PROD}^{\mathrm{dir}} + \mathrm{RES}$.

The monetary balance of payments (BP) is net specie imports *plus* the change in nonspecie net foreign assets of the authorities:

$$\mathrm{BP} = \mathrm{IMP} + \Delta\mathrm{NFA} \tag{22.3}$$

With the alternative measures of IMP, (a) $\mathrm{BP}^{\mathrm{dir}} = \mathrm{IMP}^{\mathrm{dir}} + \Delta\mathrm{NFA}$ and (b) $\mathrm{BP}^{\mathrm{ind}} = \mathrm{IMP}^{\mathrm{ind}} + \Delta\mathrm{NFA} = \mathrm{BP}^{\mathrm{dir}} + \mathrm{RES}$. The two balance-of-payments series are shown in Table 22.1, with $\mathrm{BP}^{\mathrm{dir}}$ available only from 1820. While both $\mathrm{IMP}^{\mathrm{dir}}$ and $\mathrm{PROD}^{\mathrm{dir}}$ (whence $\mathrm{IMP}^{\mathrm{ind}}$) are subject to imperfect measurement, $\mathrm{PROD}^{\mathrm{ind}}$ is a much more volatile series than $\mathrm{PROD}^{\mathrm{dir}}$, a statement not true for $\mathrm{IMP}^{\mathrm{ind}}$ versus $\mathrm{IMP}^{\mathrm{dir}}$. With $\mathrm{PROD}^{\mathrm{ind}}$ associated with $\mathrm{IMP}^{\mathrm{dir}}$, this finding suggests that $\mathrm{IMP}^{\mathrm{ind}}$ is a superior measure to $\mathrm{IMP}^{\mathrm{dir}}$ (whereas no previous work has even considered using the indirect measure of specie flow) and therefore that the "true" BP is closer to $\mathrm{BP}^{\mathrm{ind}}$ than to $\mathrm{BP}^{\mathrm{dir}}$. Empirical results are shown for both $\mathrm{BP}^{\mathrm{ind}}$ and $\mathrm{BP}^{\mathrm{dir}}$.

22.1.8 *Comparative Economic Performance of Subperiods*

The eight delimited periods of 1792–1932 are compared using three sets of performance principles: monetary-oriented criteria from a generalized exchange-market-pressure model, monetary pyramiding ratios, and measures of price and income growth and stability.

22.1.8.1 Exchange-Market-Pressure Model

The annual change in monetary-base supply ($\Delta\mathrm{BASE}_S$) is the monetary balance of payments (BP) *plus* the change in the domestic-origin component of the monetary base (ΔDOB), with the latter being the sum of net specie production (PROD) and the change in the authorities' residual contribution to the base (ΔRCON).[25] The proportionate change in BASE_S is

$$\frac{\Delta\mathrm{BASE}_S}{\mathrm{BASE}_S} = \frac{\mathrm{BP}}{\mathrm{BASE}_S} + \frac{\Delta\mathrm{DOB}}{\mathrm{BASE}_S}.$$

The most-general demand-for-base function allows only for no money illusion: $\mathrm{BASE}_D = P \bullet \mathrm{base}_d$, where BASE_D (base_d) is the nominal (real) demand for base money and P is the price level in specie (prior to 1860) or gold (from 1860) prices.[26] Taking the proportionate change in BASE_D and imposing money-market equilibrium,

$$\frac{\mathrm{BP}}{\mathrm{BASE}} = \frac{\Delta P}{P} + \frac{\Delta \mathrm{base}_d}{\mathrm{base}_d} - \frac{\Delta \mathrm{DOB}}{\mathrm{BASE}} \tag{22.4}$$

Foreign (f) money-market equilibrium:

$$\frac{\Delta \mathrm{BASE}^f}{\mathrm{BASE}^f} = \frac{\Delta P^{\mathrm{f}}}{P^{\mathrm{f}}} + \frac{\Delta \mathrm{base}_d^f}{\mathrm{base}_d^f} \tag{22.5}$$

where the foreign price level (P^{f}) is in gold currency. Taking Eq. (22.4) *minus* Eq. (22.5), rearranging terms, and adding to each side the proportionate change in the exchange rate (E, the number of units of foreign currency per dollar),

$$\mathrm{EMP} = \mathrm{DPP} + \mathrm{SB} + \mathrm{DB} \tag{22.6}$$

where

$$\mathrm{EMP} = \frac{\mathrm{BP}}{\mathrm{BASE}} + \frac{\Delta E}{E}$$

$$\mathrm{DPP} = \frac{\Delta P}{P} - \frac{\Delta P^{\mathrm{f}}}{P^{\mathrm{f}}} + \frac{\Delta E}{E}$$

$$\mathrm{SB} = \frac{\Delta \mathrm{BASE}^f}{\mathrm{BASE}^f} - \frac{\Delta \mathrm{DOB}}{\mathrm{BASE}}$$

$$\mathrm{DB} = \frac{\Delta \mathrm{base}_d}{\mathrm{base}_d} - \frac{\Delta \mathrm{base}_d^f}{\mathrm{base}_d^f}$$

Equation (22.6) divides EMP (exchange market pressure in favor of the dollar) into three components: DPP (deviation from purchasing power parity in favor of the dollar), SB (monetary supply-side nominal contribution to EMP), and DB (monetary demand-side real contribution to EMP).[27] EMP has alternatives $\mathrm{EMP}^{\mathrm{dir}}$ ($\mathrm{EMP}^{\mathrm{ind}}$), resulting from $\mathrm{BP}^{\mathrm{dir}}$ ($\mathrm{BP}^{\mathrm{Ind}}$) in its construction; similarly, it has $\mathrm{SB}^{\mathrm{dir}}$ ($\mathrm{SB}^{\mathrm{ind}}$) from $\mathrm{PROD}^{\mathrm{dir}}$ ($\mathrm{PROD}^{\mathrm{ind}}$) (via $\Delta\mathrm{DOB}$). In Eq. (22.6), $\mathrm{SB}^{\mathrm{ind}}$ ($\mathrm{SB}^{\mathrm{dir}}$) is associated with

Table 22.4 Exchange market pressure (period means—percentages)

Period	*Net specie imports (period means, percent)*			
	Algebraic value		*Absolute value*	
	Direct	*Indirect*	*Direct*	*Indirect*
1792–1810		0.83		7.09
1811–1816		7.04		8.20
1817–1838	2.06[a]	2.74	9.29[a]	6.51
1839–1846	5.26	4.89	8.69	7.76
1847–1861	−14.88	−17.00	15.93	17.89
1862–1878	−8.76	−10.26	8.86	10.27
1879–1913	0.45	0.63	1.95	2.69
1914–1932	2.08	2.07	4.42	4.87
1920–1932[b]	1.31	1.41	3.70	3.78

[a]Specie-flow calculation indirect for 1817–1819
[b]Excludes years during which London gold market was nonoperational

EMP^{dir} (EMP^{ind}). DB is computed from Eq. (22.6) residually, whence $DB^{dir} = DB^{ind}$.

Considering the left-hand side of Eq. (22.6), the magnitude of EMP measures external disturbance to the domestic economy, involving a change in the monetary base and/or the exchange rate, with adjustment and possibly associated costs to follow. Table 22.4 shows the period means of both algebraic and absolute values of EMP (in percentages).[28] Period efficiency varies inversely with the magnitude of either measure of EMP, but the tougher test is absolute value, as positive and negative figures reinforce rather than offset one another. Irrespective of the criterion and of whether EMP^{dir} or EMP^{ind} is considered, the classic gold standard (1879–1913) exhibits the greatest efficiency, with the Federal Reserve period being second (absolute-value measure).[29] Removing the years during which the London gold market was nonoperational (wherefore correction for paper-currency depreciation could not be made) enhances performance of the Federal Reserve period but insufficiently for displacement of the primacy of 1879–1913.

From the right-hand side of Eq. (22.6), $|SB + DB| = MC$ is the absolute "monetary component," or the magnitude of that part of EMP contributed by the monetary supply side and demand side offsetting or reinforcing each other. The smaller the MC (as a period mean), the more

Table 22.5 Monetary component of exchange market pressure (period means)

Period	*Net Specie Imports*			
	Absolute		*Relative*	
	Direct	*Indirect*	*Direct*	*Indirect*
1792–1810		11		64
1811–1816		6		28
1817–1838	9[a]	7	56[a]	55
1839–1846	8	8	44	47
1847–1861	18	19	69	64
1862–1878	10	11	51	55
1879–1913	2	3	36	44
1914–1932	7	7	66	66

[a]Specie-flow calculation indirect for 1817–1819

efficient the period. For MC = 0 (SB = −DB), there is perfect efficiency (complete offsetting of supply and demand), but there is no maximum value of MC. For a relative measure, suppose that |SB| and |DB| are given to the monetary standard. Then their sum is the level against which MC is measured, whence the relative monetary component RMC = 100 · MC/(|SB| + |DB|), computed as a period mean. Maximum efficiency, RMC = 0, occurs again for MC = 0, but now maximum RMC = 100 (for SB · DB > 0), involving reinforcement (or non-offsetting) of supply and demand contributions.

Table 22.5 shows the MC and RMC measures. The classic gold standard has maximum efficiency for MC and shares it with the 1811–1816 interregnum for RMC. The uniform superiority of 1879–1913 over central-banking periods is especially noteworthy.

·

22.1.8.2 *Monetary Pyramiding Ratios*

The ratio BASE/SPST measures discipline, from a specie-standard viewpoint, in restricting the monetary base. Under a pure specie standard, the ratio is unity. The Independent Treasury (1847–1861) and the preceding interregnum come closest to the ideal ratio (zero coefficient of variation around a unitary mean), with results in Table 22.6 for the mean and coefficient of variation. As would be expected, the greenback period is least

Table 22.6 Pyramiding ratios

Period	*Monetary Base to Specie Stock*[a]		*Money Income to Monetary Base*[b]	
	Mean	*Coefficient of Variation (percentage)*	*Mean*	*Coefficient of Variation (percentage)*
1792-1810[c]	1.22	6.86	21.89	21.43
1811–1816	1.25	17.80	27.01	13.40
1817–1838	1.27	11.45	23.08	13.53
1839–1846	1.06	4.69	18.72	7.84
1847–1861	1.00	10.81	15.88	11.81
1862–1878	3.72	42.77	13.48	24.06
1879–1913	2.17	14.80	10.08	6.61
1914–1932	1.65	15.31	12.64	16.60

[a]Monetary base and specie stock: end of year
[b]Monetary base average of current and previous end-of-year
[c]1791–1810 for monetary base to specie stock

disciplined, followed by the classic gold standard (for mean ratio). Paradoxically, the flexibility of the ratio may help to explain the latter-period's remarkable efficiency, in both the external economy (discussed above) and the internal economy (considered below).

Consider the further ratio $(P \cdot Y)/\overline{\text{BASE}}$, where Y is real GNP and $\overline{\text{BASE}} = (\text{BASE}_{-1} + \text{BASE})/2$. The numerator of this ratio is nominal GNP denominated in gold dollars, consistent with the expression of BASE. Then the ratio is income velocity with reference to the monetary base.[30] Period efficiency involves a low and stable velocity. Therefore, measured efficiency varies inversely with the mean and coefficient of variation of velocity. Table 22.6 shows that 1879–1913 prevails over all periods as having maximum discipline (lowest mean velocity and lowest coefficient of variation).

22.1.8.3 Price and Income Behavior

Period efficiency varies inversely with price instability and price volatility. Two price concepts are used: (1) "gold-price level," the price concept (P) in the exchange-market-pressure model, which corrects for depreciation of paper currency against gold or specie, thus placing paper standards on an equal footing with effective specie standards; and (2) "paper price level," the conventional concept, in which prices are undeflated. The indicator of price instability is mean inflation (percentage per year), computed as $100 \cdot \Delta \ln P$, while the measure of volatility is the standard deviation of trend-corrected P, $100 \cdot [\ln P - F(\ln P)]$, where F is the Hodrick-Prescott filter (smoothing parameter 100). Adjustment for trend eliminates bias in period comparisons, and for each statistic perfect efficiency involves a zero value. Results are exhibited in Table 22.7. The classic gold standard has the best performance by either criterion and for each price concept. The Federal Reserve period exhibits maximum volatility for each price level and, if truncated at 1929, maximum instability for the gold price. The 1811–1816 period has the most unstable, and the second most volatile, paper prices.

Income growth and cyclical stability are the final efficiency criteria. A fair comparison of monetary standards is enhanced by expressing growth (percentage per year) in per-capita terms: $100 \cdot \Delta \ln(YC)$, where YC is per-capita real income, but cyclical income is in overall terms and trend corrected, $100 \cdot [\ln(Y) - F(\ln Y)]$. Monetary-standard performance increases with mean growth and decreases with the standard deviation of cyclical income. As shown in Table 22.7, 1879–1913 is trivially behind the greenback period in highest mean growth but substantially superior to it in income stability. Truncating the Federal Reserve period to end

Table 22.7 Monetary-standard performance: price and income

Period	*Mean rate of growth (percentage)*			*Standard deviation*[a] *(detrended logarithm)*		
	Gold Price	*Paper Price*	*Per-Capita Income*	*Gold Price*	*Paper Price*	*Income*
1792–1810	1.26	1.26	1.04	5.09	5.24	3.46
1811–1816	2.62	4.06	0.30	6.86	9.34	3.20
1817–1838	−1.48	−1.80	0.94	6.45	6.44	3.32
1839–1846	−2.63	−2.82	0.32	5.60	5.81	3.67
1847–1861	1.19	1.19	1.50	4.15	7.73	4.18
1862–1878	0.46	0.50	1.94	7.94	8.56	4.50
1879–1913	0.21	0.19	1.93	2.66	2.77	3.44
1914–1932	1.65	1.65	−0.50	9.73	9.73	7.50
1914-1929[b]	3.47	3.47	1.65	10.08	10.08	6.60

[a]Multiplied by 100
[b]Truncated period ending with year of peak income

in 1929, the year of peak income, transforms its negative growth to only third highest, and the other central-bank periods exhibit growth below even that of the Independent Treasury System.

The classic gold standard also performs well in cyclical stability, third to the 1811–1816 and Second Bank periods. Even with the 1929 truncation, the Federal Reserve period shows maximum instability, followed by the greenback period.

22.1.8.4 Comparison with Existing Literature

The existing literature on economic performance of historical periods is quite different from the current study.[31] The usual objective is multi-country comparison of the classic gold standard and/or comparison of that era with later periods, whereas the current study is strictly concerned with the United States and over a long time span. Previous studies ignore the greenback period, rarely consider the antebellum experience, and do not delineate pre-1914 subperiods by monetary authority. Also, the existing literature disregards exchange market pressure and monetary pyramiding, considering only price and income behavior. The only findings at all comparable to the current study are Meltzer (1986) and Miron (1989), and for truncated classic gold standard and Federal Reserve periods. Meltzer's results are unfavorable to the classic gold standard, which has greater "risk" and "uncertainty" for real output, whereas Miron's findings are consistent with the current study.

22.1.9 Conclusions

Friedman and Schwartz's (1963) book is properly judged as "surely one of the most important books in economic history, and indeed, in all of economics, written in the twentieth century" (Rockoff, 2000, p. 1). The current study builds on the FS fundamental data contribution, their monetary-base series, and extends it back to 1789. The result serves as the foundation for a monetary balance-of-payments series (in two versions) over 1790–1932—a time span during which the United States was actually or potentially on a specie standard.

The years 1792–1932 are divided into eight periods (including three central-banking episodes, the Independent Treasury System, the greenback period, and the classic gold standard), and their economic performances are compared. A generalized exchange-market-pressure model naturally evolves from the data development, with the balance of payments deflated by the monetary base. Criteria of monetary-standard efficiency, developed from this model and also from monetary-pyramiding considerations, are used in addition to the usual price and income behavior. The main empirical result is the undeniable superiority of the classic gold standard (1879–1913) over central-banking episodes and all other periods.

Appendix

Data Sources and Construction of Variables

Specie Stock

1789–1859. Existing specie-stock series for the antebellum period all are generated by adding a net specie-flow series cumulatively to a base Fig.[32] The principal problem with this technique is the incomplete nature of the official specie trade series, due to (1) smuggling, (2) unavailability of a reliable return on silver imports from Mexico, (3) absence of the requirement that overland exports to Canada and Mexico be declared, and (4) lack of data on gold and silver brought in by immigrants (until the mid-1850s).[33] It follows that a superior method involves *benchmark specie-stock estimates emanating from data on specie held by the banks (including the central bank), Treasury, and public*, without the use of flow data. An annual specie-flow series then serves to interpolate between benchmark dates, with resort to linear interpolation where the series is inadequate or unavailable. The underlying assumption is that the benchmark figures are better estimates than those obtained by cumulating specie flows.

For end-of-year 1789–1806, Blodget (in *Treasury Report*, 1855, p. 51) is the specie-stock (SPST) source.[34] Using only stock data, Blodget in effect counts specie held by the banks and public (FS 1970, p. 233). Gallatin (1831, pp. 45, 49, 53–54), who was well aware of Blodget's work, provides figures for specie in banks for end-of-year 1810, 1814, 1815, 1819, and 1829. He generates the (end-of-year) 1829 specie stock explicitly as the sum of specie in banks and in the public. His technique for estimating public holdings of specie is applied here to his data for the other years, resulting in corresponding benchmark specie-stock figures.

Woodbury provides specie held by banks and by the public for various dates; his basic figure is for the end of 1833, providing another benchmark estimate.[35] The final benchmark figure is for end-of-year 1860 and is the gross specie stock, constructed as the sum of the gold stock (computed below) and subsidiary silver stock (the latter "midyear-averaged" [the average of the current and subsequent midyear] figures for 1860 and 1861, in Treasury *Report*, 1928, p. 552).

Specie-stock figures between the benchmark dates remain to be determined. The interpolative flow series (F) for 1820–1860 is constructed as the sum of net specie imports and net specie production, where net specie production is domestic production *minus* nonmonetary consumption. This consumption involves gold or silver obtained via import of coin or bullion, domestic production of bullion, or melting of coin (but not via reworking of existing nonmonetary metal) that is retained as bullion or used in arts and industry *minus* nonmonetary metal (in jewelry or other manufacture) melted down and recoined.

Net imports of gold and silver are official data, available from 1820 and customarily used by researchers.[36] However, better production and consumption series have become available since earlier research or were ignored in previous work on the specie stock. Annual gold production for 1820–1847 is the "middle" estimate of Martin (1976, pp. 446–447), with the total for 1792–1823 divided equally among the years. For 1848–1860, the source is Berry (1984, pp. 74, 76). Silver production is from Herfindahl (1966, pp. 323, 328–329).[37] Seaman (1852, pp. 258–260) is the source of nonmonetary consumption for 1820–1850.[38] Figures for 1851–1860 are obtained via linear interpolation between 1850 and 1880 (new gold and silver used in manufacturing and the arts, in *Mint Report*, 1921, pp. 62–63).[39] Gold consumption in 1880 is gold used in manufactures and the arts, total new material (Mint *Report*, 1921, p. 62). Silver consumption in 1880 is the product of ounces of silver used in manufactures and the arts, total new material (Mint *Report*, 1921, p. 63), and the price of silver (Census 1975, p. 606).

To interpolate between successive benchmark estimates, years 0 and n, let $\Delta_0^n \text{SPST} = \text{SPST}_n - \text{SPST}_0$ (with the gross specie stock used in place of SPST_{1860}, which is purely gold), and note that $\Delta_0^n \text{SPST}$ is uniformly positive. Linear interpolation is applied where F is unavailable (1807–1809, 1811–1813, 1816–1818) or $\sum_{i=1}^{n} F_{\text{i}}$ is negative (1820–1828). This leaves (n, 0) = (1833, 1829) and (1860, 1833). Intervening years $j = 1, \ldots, n - 1$ are obtained as.

$$\mathrm{SPST_j} = \mathrm{SPST_0} + \left(\frac{\sum_{i=1}^{j} F_\mathrm{i}}{\sum_{i=1}^{n} F_\mathrm{i}} \right) \cdot \Delta_0^\mathrm{n}\,\mathrm{SPST}$$ [40]

1860–1932. The gold-stock series constructed by FS (1963, p. 723; 1970, pp. 353–354) for 1867–1932 is replicated and extended back to 1860, with some different data sources. Published figures for 1860–1877 are only for midyear (June 30). Except for 1873–1879 in Mint *Report* (1941, p. 84), pre-1879 figures are not corrected for gold presumed lost by the Director of the Mint. Unrevised data for midyear 1860–1872 and end-of-year 1878 are in Treasury *Report* (1898, p. 109; 1928, p. 552).

The FS gold-loss series for midyear 1867–1873 is the difference between uncorrected (Treasury *Report*, 1928, p. 554) and corrected (Kindahl, 1961, p. 40) gold plus gold certificates in circulation. Rounded to the nearest million, the figure equals that shown by FS (1963, p. 17) for 1867 and is readily extended to 1860 via FS's linear interpolation. Subtracting gold loss from the unrevised stock, the corrected stock series is obtained for midyear 1860–1872. End-of-year specie stock for 1860–1877 is computed via midyear averaging. For 1878, it is the difference between the uncorrected stock and the average of the midyear 1878 and 1879 gold loss. For end-of-year 1879–1932, the source is Mint *Report* (1941, p. 84).

Net Specie Imports and Production

Net specie imports ($\mathrm{IMP^{dir}}$) are from Sect. 22.1.10.1 for 1820–1859, calendar-year annualized net gold imports (same source) for 1860–1932, *minus* calendar-year annualized addition to gold exports to Canada (Simon, 1960, p. 645) for 1860–1893, *plus* change in earmarked gold (Board of Governors of the Federal Reserve System [hereafter "Governors"], 1943, p. 536) for 1916–1932. Net specie production ($\mathrm{PROD^{dir}}$) is production *minus* nonmonetary consumption, where production is from Sect. 22.1.10.1 for 1820–1859, computed as described in Sect. 22.1.10.1 for 1792–1819, the 1792–1823 annual value repeated for 1789–1791, and gold production for 1860. For 1861–1900, the source for gold production is Berry (1984, p. 78); for 1901–1932, production in fine ounces (Census 1975, p. 606) is multiplied by price ($20.67 per fine ounce).

Nonmonetary consumption is from Sect. 22.1.10.1 for 1820–1859 and computed as described in Sect. 22.1.10.1 for 1811–1819. The percentage decline 1831–1821 is applied to 1801–1791 and 1811–1801 (as for 1821–1811), and linear interpolation is used for intervening values. Consumption for 1860–1932 is of gold alone. For 1860, gold consumption is estimated as the product of 1860 specie consumption and the 1880 ratio of gold consumption to gold plus silver consumption (from Sect. 22.1.10.1). Linear interpolation between 1860 and 1880 is used for the intervening years. Gold consumption for 1881–1932 is gold used in manufactures and the arts, total new material (Mint *Report*, 1921, p. 62; 1933, p. 30).

Lost and Foreign-Held Currency

Lost currency. Lost Treasury currency (including gold certificates) and Federal Reserve currency, midyear 1862–1933, is obtained as the difference between listed and loss-adjusted circulation of currency denominated up to $1000.[41] Estimated lost national bank notes are excluded during 1862–1874, via subtraction of the product of the computed lost currency and the official circulation ratio of national bank notes to the sum of national bank notes, old demand notes, U.S. notes, and gold certificates (data from Treasury *Report*, 1928, p. 554, and for old demand notes, from Treasury *Reports*, 1862–1874; the ratio is zero during 1862–1863). Midyear averaging of the resulting series yields end-of-year figures for lost currency 1862–1932.

Foreign-held currency. Countries for which dollar holdings are available are those in Europe (including Britain), Cuba, the Dominican Republic, and Honduras. The initial European stock of dollars at end-of-April 1923 (Governors 1943, p. 417) is assumed to derive from equal annual flows, beginning with a zero stock at the end of 1917.[42] The April 1923 stock is assumed to equal the measured cumulative net outflow May 1923 to June 1938; with the May-December flow included to yield end-of-year 1923 and subsequent annual net flows added to provide end-of-year 1924–1932 figures; this is Garber's (1986, pp. S150–S151) methodology.

The stock of dollars in Cuba is from Wallich (1950, pp. 320, 324–325). He provides an end-of-year series for 1920–1932. For 1912, the stock of dollars is taken as the midpoint of Wallich's range for coin plus dollar bills *minus* the midpoint of the range for coin. The dollar stock in 1912 is interpolated linearly to zero in 1897, on the assumption that dollars in Cuba reached a measurable level only with the Spanish-American War.[43] For 1918, Wallich adds a range of $10 million to $15 million to the contemporary estimate of coin plus dollar bills and interprets the contemporary-author's method as dollars constituting two-thirds of the total. Taking the midpoint of Wallich's range, the arithmetic is clear for the 1918 estimate. Linear interpolation yields figures for 1913–1917 and 1919.

The data source for dollars held in the Dominican Republic and Honduras is Mint *Reports* (1917–1931). It is assumed that an April 1917 figure for the Dominican Republic applies to end-of-year 1916. In 1905, President Theodore Roosevelt imposed a customs receivership on the country, and the dollar was adopted as the standard of value. Therefore, the dollar stock of the Dominican Republic is deemed to have increased linearly from zero in 1904 to the 1916 figure. Existing end-of-year figures are 1917, 1919–1920, 1922–1923, and 1925–1930 for the Dominican Republic and 1920–1921 and 1924 for Honduras.[44] Linear interpolation between adjacent years is applied to obtain 1918, 1921, and 1924 for the Dominican Republic and 1922–1923 for Honduras. It is assumed that 1931–1932 values for the Dominican Republic are the same as the 1930 value.

As a consequence of a rise in the price of silver above 70 cents per ounce (1917–1920 according to Census 1975, p. 606), $3 million in U.S. currency was imported into Honduras by end-of-year 1920 (Mint *Report*, 1921, pp. 154–155). This amount is allocated equally over these 4 years. The residual stock at end-of-year 1916 is an end point for linear interpolation to 1904, as for the Dominican Republic. In 1926, the government of Honduras took steps to discourage dollar circulation (Mint *Report*, 1927, p. 127). Therefore, the 1925 figure, assumed to be the same as the 1924 figure, is halved for 1926 and halved again for 1927–1932.

First Bank Variables

Specie. Specie holdings for 1792–1800 are from the Bank's (closest to) end-of-year balance sheets prepared by Wettereau (1985); for 1791, with no branches, the figure is for the Philadelphia main office alone. For 1801, 1808, and 1810, Wettereau's presentation of the November 26, 1801; February 1809; and January 15, 1811, balance sheets of Gallatin are used. For 1802, the figure of $9 million is taken, based on Gallatin's statement in November that specie holdings were more than $8 million and still increasing (Wettereau, 1937). The "alarmingly low figure" in May 1804 and February 1805 suggests an amount of $2 million for (end of) 1804, exceeded only from 1797. By May 1806, with note circulation about $5 million, "the specie problem was no longer acute, the supply on hand exceeding the total note circulation" (Wettereau 1985, p. 283), implying holdings of $6 million for end-of-year 1806. Linear interpolation between adjacent figures is used for the remaining years.

Currency. The same sources as for specie provide circulation for 1791–1801, 1808, and 1810. Figures for 1803 and 1807 are taken from House Document 27, 23rd Congress, 2nd session (hereafter "HD27 23C 2s"). Linear interpolation between adjacent figures is used for the remaining years.

Non-Treasury deposits. The same sources as for specie are used for 1791–1801, 1808, and 1810, but only total deposits are given for 1791 and 1808. To estimate non-Treasury deposits for 1791, the proportion of non-Treasury deposits for March 9, 1792, is applied. Treasury deposits at the Bank, available for 1791–1801 and 1810, are obtained for 1803–1806 from Holdsworth and Dewey (1910, p. 60) and estimated via linear interpolation of adjacent years for 1802 and 1807–1809. This permits computation of non-Treasury deposits for 1808 as a residual and for 1809 as the average of 1808 and 1810.

Assume that the modern reserve ratio, defined by the equation "reserve ratio *equals* specie holdings *divided by* the sum of currency in circulation, non-Treasury deposits, and Treasury deposits," was a meaningful statistic for this conservatively operated bank.[45] In 1802, specie holdings, at their highest level to that date, are in the same order of magnitude as in 1809; in 1803, they are very close to the 1800 and 1801 values; and in 1804, they are extremely low, taken as $2 million. Therefore, it is assumed that (1) the reserve ratio for 1802 is the same as that for 1809, (2) the reserve ratio for 1803 is the average of the 1800 and 1801 values, and (3) the reserve ratio for 1804 is the average for 1792–1796, the previous years when specie holdings were less than $2 million. For 1805–1807, the reserve ratio is linearly interpolated between 1804 and 1808. Non-Treasury deposits for 1802–1807 are then obtained via the reserve ratio-equation.

Net foreign assets. These are net assets on "foreign transactions" account *plus* holdings of foreign bills of exchange *minus* Amsterdam loan outstanding. The source is the same as for specie.[46] Except for the Amsterdam loan, foreign assets and liabilities are listed in the Bank's balance sheets only for 1792–1795. Using information in Wettereau (1937, p. 269, n. 27), a complete series of the outstanding amount of the Amsterdam loan—a relatively large item—is constructed. It is unknown whether there were other foreign items during the years for which balance sheets are not extant.

Second Bank Variables

Regarding Second Bank data, Smith (1953, p. 276) is suspicious of the much-used Tyler Report (Senate Document 17, 23rd Congress, 2nd session [hereafter "SD17 23C 2 s"]) and recommends using the Bank's actual returns whenever possible, the procedure followed here.

Specie. Consulting the end-of-year returns printed in various congressional documents, Bank specie is obtained for 1821–1823 and 1825–1838.[47] For the remaining years, resort must be had to Tyler's data (in SD128 25C 2 s). There are no data for end-of-year 1817, so linear interpolation is applied to the figures for September 1817 and February 1818.

Currency. Circulation for 1817–1820, 1824, and 1832–1838 is from the same sources as for specie. For the remaining years, SD128 25C 2 s is used; because the pre-1832 returns show only notes issued, notes on hand and in transit must be deducted to derive circulation.

Non-Treasury deposits. Same sources as specie are used.

Net foreign assets. Holdings of foreign bills (or net foreign exchange) *plus* amount due from European bankers *minus* amount due to European bankers are used. The sources are the same as for specie, except for 1837. For that year, linear interpolation is applied to the figures for December 1, 1837, and February 1, 1838 (from actual returns in SD128 25C 2 s).

Federal Reserve Variables

Specie. Gold in Federal Reserve banks is from Mint *Report* (1941, p. 84).

Currency. Federal Reserve notes and Federal Reserve banknotes in official circulation are from Governors (1943, pp. 409–412), the FS source.

Domestic non-Treasury deposits. Domestic bank deposits at Federal Reserve Banks (FS 1963, pp. 737–740) are used.

Net foreign assets (other than foreign-held currency). Holdings of foreign bills *plus* amount due from foreign banks *minus* foreign deposits at Federal Reserve banks (Governors 1943, pp. 330–332) are used.

Lost and foreign-held currency. Consider the FS monetary-base series for end-of-November 1874–1906, end-of-February 1874–1907, and end-of-year 1907–1932. For comparability with the new monetary base, state bank notes are excluded. The FS source for state bank notes is Governors (1943, p. 408), which shows $1 million in circulation June 1873 to June 1878 and then zero. So $1 million is deducted from the November figures 1874–1877 and February figures 1874–1878. Linear interpolation, as in note 23, is applied to obtain an end-of-year series 1874–1932 (hereafter "the adjusted FS series"). Estimated Treasury and Federal Reserve currency in official circulation 1914–1932 is obtained by subtracting (1) specie in circulation (specie stock [from Sect. 22.1.10.1] *minus* Federal Reserve gold *minus* Treasury gross specie [from Sect. 22.1.10.7]), (2) nongold coin in circulation (from Sect. 22.1.10.7), and (3) domestic bank deposits at Federal Reserve banks. For 1914–1932, one computes the annual ratio of Federal Reserve currency to the estimated sum of Treasury and Federal Reserve currency in official circulation. This ratio multiplies "lost currency net of lost currency in 1913" (Sect. 22.1.10.3) to yield lost Federal Reserve currency. It multiplies "foreign-held currency net of foreign-held currency in 1913" to estimate foreign-held Federal Reserve currency.

Treasury Variables

Specie. The position that Treasury specie is zero for 1789–1835 is irrefutable (Treasury *Report*, 1915, p. 374; FS 1970, pp. 245–246). It is the sense of FS (1970, p. 248) that this is true also for 1836–1846, which appears to be the position of Taus (1943, pp. 222–224), except for the period of the first Independent Treasury. Therefore, through 1846, Treasury specie is taken as zero except for end-of-year 1840, where gold is assumed to constitute half of the balances in Treasury offices (from Treasury *Report*, 1915, p. 374). For end-of-year 1847–1859, the Macesich data published in FS (1970, pp. 222–224) are used.

For 1860–1861 and 1864–1872, midyear data are computed as the (uncorrected) stock of gold coin and bullion *minus* circulation of gold coin from Treasury *Report* (1928, pp. 552–554). The ratio of Treasury gold to the corrected gold stock for midyear 1861 and 1864 is interpolated linearly to obtain midyear ratios for 1862–1863, which in turn multiply the corrected stock to estimate midyear Treasury gold for these years. For 1873–1878, midyear Treasury gold is from Mint *Report* (1941, p. 84). Midyear averaging yields Treasury gold end-of-year 1860–1877. The figure for end-of-year 1878 is in Treasury *Report* (1898, p. 59). Treasury gold for end-of-year 1879–1932 is from Mint *Report* (1941, p. 84).

Gold certificates. Gold certificates were first issued in 1865, but in a trivial amount (Bayley, 1881, p. 162). For 1866–1877, midyear averaging is performed on official circulation data in Treasury *Report* (1928, p. 554). End-of-year data 1878–1932 are from Governors (1943, pp. 409–412) and Treasury *Report* (1898, pp. 131–132; 1903, pp. 219–220; 1909, pp. 204–208; 1915, pp. 351–354).

Nongold coin in circulation. The sum of silver dollars, subsidiary silver coin, and minor coin is used. Silver dollars were not in circulation during 1860–1877. However, following FS (1963, pp. 113–114, n. 37; 723), the 1877 figure is taken as $6 million, representing circulation of trade dollars. Standard silver dollars in circulation are available end-of-year 1878–1932 (Governors 1943, pp. 409–412; Treasury *Report*, 1898, pp. 124–126; 1903, pp. 215–216; 1909, pp. 194–198; 1915, pp. 343–346).

Considering subsidiary silver coin, for 1860–1873 midyear averaging is applied to data from the ultimate source: Treasury *Report* (1928, pp. 552, 554). For midyear 1860–1863, only the stock figure is available, and the 1864 circulation/stock figure is used to estimate circulation. For 1874–1877, midyear averaging is applied to figures in Governors (1943, p. 408), the FS data source (containing fewer significant digits than Treasury *Report*, 1928). For 1878–1932, the sources are the same as for silver dollars. Following FS (1963, p. 723; 1970, p. 355), the overestimate deducted in mid-1910 by the Director of the Mint is apportioned linearly over 1881–1909.

Minor coin in circulation is available midyear 1900–1914 and end-of-year 1914–1932 (Governors 1943, pp. 408–412). Midyear averaging is applied to the former.

Currency, 1812–1873. For Treasury notes (1812–1873), outstanding issues are taken from public-debt statements: end-of-year 1812–1843 from Elliot (1845, pp. 906–917) and various end-of-quarter dates 1844–1874 from Treasury *Reports*. The latter figures are converted to end of year via (1) addition of quarterly issues *minus* redemptions from Bayley (1881) or (2) linear interpolation, used where the two adjacent known figures both are below $1 million (in fact, below $0.65 million) *and* issues are zero during the interpolation period.[48] For old demand notes, Treasury notes of 1863, compound-interest notes, and 3% certificates (1861–1873), end-of-year figures for the initial year are the issues (with redemptions again zero) (Bayley, 1881, pp. 153, 161–163, 169). Then midyear-averaged figures in outstanding public-debt statements (Treasury *Reports*, 1862–1874) are taken.[49] For U.S. notes (greenbacks, 1862–1873), official circulation is obtained via midyear averaging of data in Treasury *Report* (1928, p. 554). For fractional currency (1862–1873), the same applies, with two amendments. First, following FS (1963, p. 724; 1970, pp. 354–355), all but $1 million of outstanding fractional currency in midyear 1878 is assumed lost, distributed linearly over 1863–1878, and deducted from the official data. Second, the initial (end-of-1862) figure is total issues during 1862 (there were no redemptions [Bayley, 1881, pp. 159–160]) multiplied by the circulation/stock ratio of midyear 1863, with the estimated loss subtracted.

The sum of all the above components *plus* gold certificates in official circulation (from 1866) *minus* lost currency yields Treasury currency (with no gold certificates) in domestic circulation 1812–1865 but inclusive of gold certificates 1866–1873.

Currency, 1874–1932. The adjusted FS series *minus* specie in circulation *minus* lost currency and foreign-held currency *minus* nongold coin in circulation yields Treasury currency, inclusive of gold certificates, in domestic circulation 1874–1913. The series for 1914–1932 is obtained as this result *minus* Federal Reserve currency in official circulation *plus* lost Federal Reserve currency *plus* foreign-held Federal Reserve currency *minus* domestic bank deposits at Federal Reserve Banks.

Currency, exclusive of gold certificates, in domestic circulation, 1866–1932. This is obtained by *subtracting* gold certificates in official circulation and *adding* lost gold certificates. Lost certificates are the product of (1) the ratio of official circulation of gold certificates to that of old demand notes, U.S. notes, national bank notes (from 1874), silver certificates, Treasury notes of 1890, and gold certificates, and (2) lost Treasury currency inclusive of gold certificates, net of lost currency in 1865.

For 1866–1873, the denominator of the ratio consists of old demand notes, U.S. notes, and gold certificates. For 1874–1913, the denominator is estimated as the adjusted FS monetary base *minus* specie in circulation *minus* nongold coin in circulation. For 1914–1932, Federal Reserve currency in official circulation and domestic bank deposits at Federal Reserve Banks (Sect. 22.1.10.6) are also subtracted. Lost Treasury currency is total lost currency (Sect. 22.1.10.3) *minus* lost Federal Reserve currency (Sect. 22.1.10.6).

Prices

Specie price of currency. For the central bank, this variable is unity except for May 1837 to August 1838, when the Second Bank suspended specie payments. The percentage premium (PR) on American gold at Philadelphia for end-of-year 1837 is linearly interpolated between December 9, 1837, and January 6, 1838, observations (SD457 25C 2 s). The specie price of currency is then $1/(1 + PR/100) = 0.9609$. Non-unity specie price of currency for the Treasury is the gold price of greenbacks for the last market day of the year, 1861–1878.[50]

Price level. The paper price level is measured by the GNP deflator. For 1792–1869, the source is Berry (1988, p. 21), ratio-linked in 1869 to the series for 1869–1932 in Balke and Gordon (1989, pp. 84–85) and Department of Commerce (1986, pp. 1, 6). The gold price level (P) is the product of the paper price level and the specie price of currency (for the full year rather than end of year), with par equaling unity.

The annual specie price of currency for the antebellum period is derived as follows. The Berry deflator is based on the Hoover and Taylor (1959) composite index of wholesale price indexes in various cities. Let PCUR^i denote the specie price of currency in city i. The weighting pattern of the Hoover-Taylor index (differing for 1800–1815 from 1816–1861) is applied to the data-available PCUR^i for periods during which at least one city is on a paper standard ($\text{PCUR}^i < 1$), based on information in Officer (1996, pp. 16–17) and Berry (1943). Thus, the specie price of currency is a weighted average of PCUR^i for New York and Philadelphia (1814–1817); Cincinnati and the other cities (for which $\text{PCUR}^i = 1$) (1818–1820); New York, Philadelphia, and Cincinnati (1837–1842—but $\text{PCUR}^i = 1$ for New York 1839–1842); and New York, Cincinnati, and New Orleans (for which $\text{PCUR}^i = 1$) (1857).[51] For 1862–1878, the specie price of currency is the gold price of greenbacks (Mitchell, 1908, p. 4).

Income

The logical source for income (Y) 1792–1869 is the real-GNP series of Berry (1988, pp. 18–20), consistent with construction of the price level. However, Berry's technique is subject to legitimate criticism for the antebellum period.[52] Fortunately, the limitations of Berry's series are overcome via the broad-concept real-GDP data of Weiss (1992, pp. 31–32). The Weiss figures, developed for nine antebellum benchmark years (1793, 1800, 1807, 1810, 1820, … ,1860), are on a per-capita basis. Multiplication by population (Census 1975, p. 8) yields YW, the Weiss GDP series.

Denoting the Berry series as YB, a revised series (YR) is derived as follows: (1) For 1860–1869, YR = YB. (2) Running t from 1850 back in time over the benchmark years (with successive such years separated by m calendar years),

$$\mathrm{YR}_t = \frac{\mathrm{YW}_t}{\mathrm{YW}_{t+m}} \cdot \mathrm{YR}_{t+m}$$

(3)

$$\mathrm{YR}_{1792} = \frac{\mathrm{YR}_{1793}}{\mathrm{YB}_{1793}} \cdot \mathrm{YB}_{1792}.$$

(4)

$$\text{Let } f = \frac{\left(\frac{\mathrm{YR}_{t+m}}{\mathrm{YB}_{t+m}} - \frac{\mathrm{YR}_t}{\mathrm{YB}_t}\right)}{m}.$$

Then $\mathrm{YR}_{t+n} = ((\mathrm{YR}_t/\mathrm{YB}_t) + n \cdot f) \cdot \mathrm{YB}_{t+n}$, $t = 1850, 1840, \ldots$, $1 \leq n < m$. The source of income 1869–1932 is the same as for the paper price level. Balke and Gordon (1989) take care to express real GNP consistent with the national accounts (Department of Commerce 1986) denomination in 1982 constant dollars, whence the price level equals 100 for that year.[53] The revised series for 1792–1869 is ratio-linked to the Balke-Gordon series in 1869. Per-capita income is the ratio of real income to population.

Foreign Variables

The foreign variables are index numbers: Britain (converted to 1913 = 1) 1791–1913, an index of Britain (0.5778 weight) and Canada (0.4222 weight) (both converted to 1913 = 1) 1913–1932. Weights are proportional to share of U.S. exports and imports during 1913–1932 (Census 1975, pp. 903–906).

Exchange rate (E). For 1791–1913, the exchange rate is based on the annual average of the quarterly sight-bill equivalent exchange rate (dollars per pound) corrected for paper-currency depreciation (obtained by reversing the procedures in Officer 1996, pp. 54–55, 64–97). The inverse of this series (whence pounds per dollar) is taken and expressed as 1913 = 1.

For 1913–1932, annual cable exchange rates (DP = dollars per pound and DC = dollars per Canadian dollar) are from Governors (1943, pp. 665, 681). The United Kingdom was on a paper standard from August 1914 to April 27, 1925, and again from September 20, 1931, as was Canada to June 30, 1926, and from January 1929. It may be noted that dollar/sterling and Canadian-dollar/sterling parity was 4.8665635, with Canadian-dollar/U.S.-dollar parity at unity. The London gold market was closed during the paper standard until September 1919. So the exchange rates are corrected for paper-currency depreciation 1919–1925 and 1931–1932 as follows.

Letting PGL denote the currency price of gold in London (the ratio of the market price of gold [from Shrigley 1935, p. 92] to the mint-parity price of gold), PRP = (1/PGL − 1) is the proportionate premium of the pound over gold (with the pound at a discount, PRP is negative). The corrected dollar/pound exchange rate is DPC = DP − 4.8665635 · PRP. Letting CP denote the Canadian-dollar/pound cable exchange rate (from Leacy 1983, series J563), the proportionate premium of the Canadian dollar with respect to gold is PRC = (4.8665635/CP) · (1/PGL) − 1, and the corrected dollar/Canadian-dollar exchange rate is DCC = DC − PRC. The inverses of DPC and DCC are then expressed in index-number form.

Price level (P^f). Considering Britain for 1790–1830, the Gayer, Rostow, and Schwartz price index (in Mitchell 1988, p. 721) is ratio-linked to the GDP deflator for 1830–1932. The latter is constructed as the ratio of current-price to constant-price GDP, with the numerator and denominator each obtained by ratio-linking earlier to the first year of later component series: Feinstein (in Mitchell 1988, pp. 831–838) expenditure (1830–1854) and "compromise" (1855–1869, 1913–1919) estimates of GDP at factor cost, Solomou and Weale (1991, p. 60; 1996, pp. 110–113) "balanced" estimate of GDP (1870–1912 and 1920–1932). The Canadian GNP deflator is from Urquhart (1993, p. 25) 1913–1926, ratio-linked to Statistics Canada (Leacy 1983, series K172) 1926–1932.

For each country, the gold price level is the product of the paper price level and the gold price of currency. The currency price of gold in London (PGL) must be extended to 1797–1821, the Bank Restriction Period of the paper pound. Quarterly averages of the price of bar gold are computed from weekly observations in *Report* ([1819] 1968, pp. 335–354) for 1797–1818 and *Report* ([1832] 1968, pp. 98–100) for 1819–1821. Annual averages of the available quarterly observations are taken, and linear interpolation is applied for missing years (1800–1803 and 1806–1809).[54] The gold price of currency for Britain is 1/PGL, while for Canada it is (PRC + 1).

Monetary base (BASE^f). The Canadian monetary base (1913–1932), from Metcalf, Redish, and Shearer (1996), is conceptually equivalent to the FS base and the new monetary base. The British series (BASE^B), developed here, differs in including only domestic bank deposits (hereafter "bankers' balances") at the Bank of England, excluding other non-central-government deposits, because the latter cannot be separated from foreign deposits.

For 1791–1869, BASE^B is the sum of coin in circulation (CC), Bank of England notes in circulation (BN), Scottish and Irish banks notes in circulation *less* coin held (SIN) (from 1845, pursuant to the Bankers' Acts [Scotland and Ireland] of that year), and bankers' balances (BB). The sources for SIN are Report ([1857] 1969) for (last date in year) 1845–1856 and *The Economist* (4-week average ending date closest to year-end) 1857–1869. For 1791–1867, CC = SP − BAC, where SP is the specie stock and BAC is the Bank of England coin and bullion. CC 1868–1869 is midyear-averaged figures of Capie and Webber (1985, p. 198). For 1844–1869, BAC and BN (constructed as notes issued *minus* notes in Banking Department) are from *The Economist*, closest return to end of year. Prior to 1844, the preferred source of any Bank series is Bank of England *Quarterly Bulletin* (June 1967, Appendix [hereafter QB]). Other series for BAC and BN are in *Reports* ([1840, 1841, 1848] 1968).

Let QBF denote the QB end-of-February series, (RF, RN, RD) the corresponding *Report* series for end of (February, November, December), and the subscript 1 the series forwarded one year. Formula A is $(QBF_1/RF_1) \cdot RD$, formula B differs in linearly interpolating RD as $(2 \cdot RN + RF_1)/3$, and formula C is $(2 \cdot QBF + 10 \cdot QBF_1)/12$. Formula A is used to estimate BN 1792–1797, BN 1815–1843, and BAC

1832–1843; formula B to estimate BN 1791, BN 1798–1814, and BAC 1816–1831; formula C to estimate BAC 1791–1815.

Benchmark year-end dates for SP are 1790 (the 1800 figure *minus* 10 times average annual net imports of specie 1791–1800 [from Brezis 1995, p. 51]); 1800, 1830, and 1860 (from Feinstein 1988, p. 397); and 1868 (composed as CC + BAC). Net imports of specie (F^{B}) are from Brezis 1791–1800, computed as $\left(\mathrm{SP}_{1830} - \mathrm{SP}_{1800} - \sum_{1816}^{1830} F^{\mathrm{B}}\right)/15$ for 1801–1815, and from Imlah (1958, pp. 70–72), changing sign of his net-exports series, for 1816–1868. The interpolative technique for SPST is then applied to SP, with F^{B} the interpolative series.[55]

For 1791–1818, BB is constructed as $U \cdot V \cdot \mathrm{TD}$, where TD is total deposits, V is the estimated ratio of private (non-central government) deposits to total deposits, and U is the estimated ratio of bankers' balances to private deposits. Data are from QB and *Report* ([1832] 1968). TD is obtained by applying formula (C and B) to (1790–1814 and 1815–1818). Considering the numerator and denominator of V: for 1791–1806, they are the annual average of 1807; for 1807–1813, they are the annual average of the current year plus the annual average of the subsequent year; for 1814, they pertain to February 1815; for 1815–1818, they are linearly interpolated as for RD in formula B. For 1791–1818, $U = \mathrm{BB}_{1819}/(V_{1819} \cdot \mathrm{TD}_{1819})$.

For 1819–1869, BB is estimated as $(\mathrm{BB}_{1870}/\mathrm{BBH}_{1870}) \cdot \mathrm{BBH}$, where BBH is bankers' balances at the Bank head office (from QB) and BB (from Capie and Webber, 1985, p. 409) also includes balances at branches.[56]

For 1870–1932, BASE^{B} is obtained from the Capie and Webber (1985, pp. 54–57) end-of-year series by *adding* Bank of England Banking Department coin (last reporting date in December, from *The Economist*) and *subtracting* Banking Department notes and coin (Capie and Webber, 1985, pp. 409–420).[57]

Notes

1. Rutner, who has performed the most thorough investigation of the central-bank status of the Second Bank, states, "The ultimate criterion by which the BUS could be a central bank and which would make it unique is simply this: did other economic actors (i.e., banks and individuals) consider BUS monetary liabilities a form of reserve currency?" (Rutner 1974, p. 121). He answers strongly in the affirmative (see below).

2. See Hammond (1957, p. 403), Warburton (1962, p. 67), Fenstermaker (1965, p. 69), Rockoff (1971, p. 456), Rutner (1974, pp. 23, 27, 143–144), Timberlake (1993, p. 241), and Highfield, O'Hara, and Smith (1996, p. 483). However, Temin treats the Second Bank as a commercial bank, albeit an important one.
3. The fact that they were not uniformly so included is not a "puzzle" (Rutner's term), for (1) Bank notes were not a legal reserve and (2) there was no minimum reserve requirement.
4. By contrast, Fenstermaker and Filer (1986) find that the Banks of the United States did not affect the behavior of New England state banks, but they view this result as purely regional.
5. Rutner (1974) observes that even "in the Panic of 1837…there appears to be fairly strong evidence to suggest that the BUS monetary obligations were considered a form of reserve currency and hence in this sense the BUS was a central bank" (p. 145).
6. Within a few months of beginning operations at Philadelphia (the head office), each Bank established branches in Baltimore, Boston, Charleston, and New York (plus 13 other locations, on the part of the Second Bank). Ultimately, the First Bank had eight branches, and the Second Bank had a maximum of 26 at one time. After the Second Bank became a Pennsylvania state bank, it continued to operate nationally by converting its branches to agencies.
7. The monetary-control argument is best made for the First Bank by Hammond (1957, pp. 198–199) and Perkins (1994, p. 249), and for the Second Bank by Temin (1969, pp. 49–53) and Timberlake (1993, p. 241). Rockoff (1971, pp. 456–457) observes that the Second Bank continued this form of monetary regulation even after it became a Pennsylvania state bank.
8. This comparison, made by historians for the Second Bank (e.g., Shultz and Caine 1937, p. 211; Smith 1953, p. 236; Studenski and Krooss 1963, p. 87), again can be extended to the First Bank.
9. The Bordo-Kydland-Rockoff thesis suggests that it is a mistake to view the greenback period (or any suspension of specie payments) as uniformly involving the weakest adherence to a metallic standard and to view the classic gold standard as uniformly involving the strongest. In fact, for most of the last decade of the greenback period, there was strong expectation of a return to the former gold standard (Bordo and Kydland 1995, pp. 451–452), and for much of the early and mid-1890s, there was a high objective and subjective probability of U.S. abandonment of the gold standard (FS, 1963, pp. 104–113).
10. It is arguable that the First and Second Banks gained their central-bank status only gradually when the institution came into existence and lost it similarly when the Bank was on its way out. In this vein, Rutner

(1974, p. 125) asks, "Did BUS [Second Bank] monetary liabilities lose their 'high- poweredness' in a continuous or discontinuous manner?" He includes Second Bank notes and deposits in the monetary base until the very end of the Bank's existence in early 1842, but he also shows an alternative series excluding the Bank's liabilities from the base. The ideal solution might be to assign weights to the Banks' liabilities increasing from zero to unity at the beginning, decreasing from unity to zero at the end, if only the weighting patterns were known. The current study, in effect, allocates a weight of unity to Second Bank liabilities until the end of 1838 and a zero weight from the end of 1839.

11. Because the First and Second Banks did not generally behave as lenders of last resort, they were "outside" agents only in the sense of having their liabilities serve as components of the monetary base. However, it is also true that the performance of the Federal Reserve System as a lender of last resort during the early 1930s was "little more than lip service" (FS, 1963, p. 395).
12. Prior to 1860, nongold coin in circulation is included in the specie stock.
13. Currency is at par when measurable amounts are held by foreigners. So there is no conversion process for foreign-held currency.
14. Throughout this study, beginning-of-year data are considered end-of-previous-year data.
15. Yet there remains a conceptual problem. The greenback price of gold is highly correlated with the price level, and depreciated monies constitute 69% of the base during the greenback period (see Tables 22.2 and 22.3). So the gold-denominated base is roughly the real base for this period. Certainly, one would not apply this procedure after 1932, and especially after 1972, when the paper-dollar nominal monetary base (constituting the entire base) would be deflated by a volatile price of gold. So legitimate comparisons between the *greenback period* new monetary base and the post-1932 base could not be readily made. Also, in the long run, the resulting new base might approximate the real base, and a nominal base does not remain for analysis. I am indebted to an anonymous referee for raising several important issues, including this one.
16. In principle, as a compromise between the two positions, national bank notes could enter the monetary base prior to 1874 but with a weight below unity.
17. Notwithstanding the Mint Act of June 28, 1834, which undervalued silver relative to gold, there is evidence that "silver coins remained in common use in the United States until some time after the discovery of gold in California [in 1848]" (Berry, 1943, p. 488). In a similar vein, Martin (1973, p. 825) shows that "de facto bimetallism…persisted to mid-century." It appears that the turning point was the Subsidiary Coinage Act of February

21, 1853 (Officer 1996, p. 20), but Berry observes that as late as 1857, silver (along with gold) coin was advertised at a premium.

18. Smith (1953, p. 182) refers to "the post notes of ill repute" and observes that "the amount of these issues was a reliable index of the degree of financial emergency within the Bank."
19. The FS (1963, p. 25, n. 10) claim that their currency figures include 3% certificates is false; see note 49. However, FS are followed in their *exclusion* from the monetary base of other interest-bearing currency issued during the Civil War. Recent assessments of the "moneyness" of various forms of interest-bearing Civil War currency are in Gherity (1993) and Woodward (1995). The evidence is mixed and intertwined with the definition of moneyness.
20. Carothers (1930, pp. 170–185, 241–261) provides the best history of these remarkable currencies. He observes that "these glue-coated bits of paper [postage currency] were the worst form of currency ever used by a civilized people" and, quoting Knox, that fractional currency "wore out rapidly and became ragged and filthy."
21. See FS (1963, pp. 207, 257, n. 40), Rutner (1974, pp. 248–253), and Sylla (1982, pp. 31–33).
22. Also, none of the authors includes the Civil War years, and only Bordo includes the antebellum period. Furthermore, FS provide charts rather than figures; Bordo and Cagan deal only with *changes* in the base; and, like Temin, Bordo defines the antebellum monetary base as composed only of specie.
23. Temin's (1969, pp. 186–187) series is at end of fiscal year (September 30, 1820–1842, and June 30, 1843–1858). Rutner's series (not seasonally adjusted, with Second Bank a central bank) is selected for compatibility with the new base. It has year-end data points except for 1835, 1840, and 1843–1846. FS (1963, pp. 800–804; 1970, pp. 344–350) provide data for end-of-November 1867–1906, end-of-February 1867–1907, and end-of-year 1907–1932; the November–February figures serve as interpolative points for year-end Figs. 1867–1906.
24. The formula to calculate annual average percentage change in X is $100 \cdot \ln(X_{t+n}/X_t)/n$, where t is the initial year and $t+n$ is the final year.
25. Taking the first-difference of Eq. (22.1) and incorporating Eqs. (22.2) and (22.3) yields $\Delta \mathrm{BASE}_S = \mathrm{BP} + \Delta \mathrm{DOB}$.
26. Throughout the model, price levels and the exchange rate are corrected for paper-currency depreciation, in conformity with the monetary base expressed in gold dollars.
27. The exchange-market-pressure model, of which Eq. (22.6) is a generalization, has been criticized by Weymark (1995, 1997a, 1997b, 1998). She argues that EMP is the simple sum of the change in official reserves (here,

balance of payments) and exchange-rate components only under restrictive assumptions: (1) purely monetary model, (2) small open economy, (3) exogenous exchange-rate intervention, and (4) exogenous change in domestic credit (a component of ΔDOB in the current model). In response, first, Weymark extends the monetary model by incorporating aggregate demand and supply (and nontraded goods), but at the cost of complexity. The monetary approach is readily operational and melds well with the monetary balance of payments. Second, introducing foreign-country exchange-market intervention into EMP is unsuitable for the purpose at hand, in which EMP is constructed to impinge directly on the domestic economy. Therefore, the totality of the foreign-base term may be placed in SB. Third, under a specie standard, exchange rates are kept within the gold-point spread typically by passive specie-transactions behavior of the authorities. Fourth, providing that the authorities respond to EMP itself or that they sterilize gold flows, the definition of EMP remains valid even with endogenous change in domestic credit. In fact, the Second Bank under Biddle altered domestic credit in response to both specie-flow and exchange-rate change (Redlich, 1968, pp. 125, 134), and the Federal Reserve System sterilized gold flows for much of the 1920s and into the second half of 1931 (FS 1963, pp. 279–287, 297, 360–361, 396–399).

28. In the construction of variables, the proportionate change in X is $\Delta X/X_{-1}$.
29. It is also true that 1879–1913 has the smallest magnitude of the absolute value of every component of EMP, no matter how composed: BP^{dir}/BASE, BP^{ind}/BASE, ΔE/E, DPP, DB, SB^{dir}, and SB^{ind}.
30. This statement is valid only if P and Y are defined so that (1) $P = 1$ in the national-accounts base year and (2) the unit of measurement of Y is the same as that for BASE. Otherwise, the ratio is income velocity only up to a multiplicative constant. P and Y are constructed to make the statement true.
31. Nearly all of the studies are listed in Bordo and Schwartz (1999). See also Basu and Taylor (1999).
32. This is the technique of Temin (1969, pp. 185–189) and Rutner (1974, pp. 205–216) as well as that of Seaman (1852, pp. 257–260); Secretary of the Treasury, *Annual Report* (hereafter "Treasury *Report*") (1855, p. 71); and Warburton (reported in FS, 1970, p. 227).
33. See Treasury *Report* (1854, p. 281; 1855, p. 71) and Simon (1960, pp. 631–632, 644).
34. Blodget's series, for 1790–1807, is dated end-of-year by FS (1970, pp. 216–219), but it is interpreted as beginning-of-year (end of the previous year) by Temin (1969, p. 185), and by FS (1970, p. 244, n.

16) themselves, via the dating of the table in Treasury Department (1915, p. 45), which includes the Blodget figure for 1800.

35. Rutner (1974, pp. 205–207) believes that this is Woodbury's basic figure, and FS (1970, p. 227) provide evidence that it is indeed so, but it is uncertain whether the estimate is derived purely from stock data. Woodbury describes his numbers only as "prepared partly from actual returns, and partly from estimates" (Elliot, 1845, pp. 941–942).
36. The source is Census (1949, pp. 243–245), with "calendar-year annualization" of figures for other than calendar years. For example, data for years ending June 30 (September 30) are allocated 50% (75%) to the current (the remainder to the previous) year.
37. Linear interpolation is used between benchmark dates, and a half-year of operation is assumed for the initial year (1836) of the only significant silver mine. Data are converted from physical output to value via multiplication by the New York price of silver (1836–1849 from Director of the Mint, *Annual Report* [hereafter "Mint *Report*"], 1910, p. 99, with price computed as the ratio of value to output; 1850–1860 from Census, 1975, p. 606). Rutner uses Herfindahl's silver (and gold) data, but only for 1834–1849.
38. Calendar-year annualization is applied as warranted. Temin makes no allowance for nonmonetary consumption. Rutner (1974) and Shetler (1973) do not employ pre-1880 consumption data for their antebellum estimates (and Rutner errs in including reworked metal), but Seaman shows an appreciation of the concept of nonmonetary consumption that vindicates his numbers. Seaman's figures for 1821–1846 are net of domestic production and require restoration to gross level. From the text, it may be inferred that he takes production as (1) essentially zero for 1821–1823, (2) deposits of domestic gold production at the mints for 1824–1829, and (3) $500,000 for 1830–1846. The figure for 1820 is obtained by assuming that the percentage decline in consumption from 1821 to 1811 was the same as that from 1831 to 1821, and applying linear interpolation.
39. This is an important antebellum interpolation. It is prudent to check whether log-linear rather than linear interpolation makes a difference. The Theil inequality coefficient between the alternative interpolative series and the actual series F is 0.0086, with zero being a perfect fit.
40. Note that linear interpolation involves (j/n) in place of the bracketed term.
41. The source is Laurent (1974, p. 221). It is reasonable to assume that large-denomination notes would be guarded most carefully.
42. Garber (1986, pp. S140–S141, S150) provides evidence that "prior to World War I little U.S. currency was held in Europe." It is unlikely that this situation changed until some time after American Expeditionary Forces arrived in France in June 1917.

43. This is quantitatively the most important linear interpolation for foreign-held dollars. Log-linearity would change the monetary base in any year by less than \$6 million, less than one-fourth of 1%.
44. A Honduras figure for 1922 is unreasonably low and so is disregarded.
45. This is the view of Perkins (1994, p. 248), who computes the ratio for various years.
46. The exchange rate to convert foreign bills in 1793 from guilders to dollars is in Wettereau (1985, p. 87).
47. The sources are HD52 17C 1 s (1821), HD78 18C 1 s (1822–1823), HD105 19C 1 s (1825), *American State Papers: Finance* 766 19C 2 s (1826), HD100 20C 1 s (1827), HD93 20C 2 s (1828), HD63 21C 2 s (1829–1830), HD523 23C 1 s (1831–1833), SD128 25C 2 s (actual returns) (1834, 1836), SD312 24C 1 s (1835), SD 471 25C 2 s (condensed return) (1837), and HD172 26C 1 s (1838). Smith did not locate returns prior to 1825.
48. There are discrepancies between Bayley's flow data and the change in amount outstanding obtained by first-differencing the public-debt series, but the divergence is of importance only for small changes in amount outstanding. For possible reasons for the discrepancy, see Treasury *Report* (1846, p. 29) and Rutner (1974, p. 253). Bayley's figures are probably superior to the Treasury flow data—the latter used by Rutner—because Bayley accounts for and corrects anomalies in the Treasury data. Also, Rutner obtains his outstanding-notes series by continuously cumulating sales *minus* retirements, a technique that fails to take advantage of the (presumed definitive) public-debt statements.
49. In using the "other U.S. currency" series rather than consulting the public-debt statements, FS commit actual or potential errors. First, "other U.S. currency," as found in Treasury *Report* (1928, p. 552), equals the sum of outstanding old demand notes, Treasury notes of 1863, and compound-interest notes; 3% certificates are excluded. Second, for midyear 1863, "other U.S. currency" is overstated by including (and thus double-counting) the stock of U.S. notes issued under the Act of March 3, 1863. At \$89.879 million, the error is substantial—20% of the 1862–1863 average monetary base. By 1874, 3% certificates outstanding are nearly zero, and there is only a trivial difference between "other U.S. currency" and the sum of the components in the public-debt statement.
50. The average of the high and low price for the day is taken, from Mitchell (1908, pp. 288–338). For 1861, the January 1, 1862, figure is used.
51. Sources of $PCUR^i$ are Gallatin (1831, p. 106) for 1814–1817, Warren and Pearson (1935, p. 154) for New York 1837–1838 and 1857, Officer (1996, p. 78) for Philadelphia 1837–1842, and Berry (1943, pp. 386–389, 398, 462, 590–591) for Cincinnati. Averages of monthly or quarterly values, often of the percentage specie premium, are taken. (Where there is

a monthly range, the midpoint is used.) If the annual value of the specie premium is PR^i, then $PCUR^i = 1/(1 + PR^i)$. Berry provides no data for June 1839 to March 1840, but the specie premium is clearly zero for June–September. He notes that the specie premium increased to about 8% within 5 months of the October 1839 suspension. It is assumed that the premium increased linearly from zero just prior to that suspension to 8% in March 1840, and a suitable weighted average of zero and the interpolated value is computed for October 1839.

52. As noted by Engerman and Gallman (1982, pp. 5, 15–16), the extrapolator series are few in number, the GNP concept excludes home production, and the extrapolations are based on a statistical model devoid of economic content.
53. Balke and Gordon (1989, p. 40) argue convincingly for their own superiority over the competing Romer (1989, pp. 22–23) series. Dividing P by 100 and expressing Y in millions rather than billions of dollars satisfies (1) and (2) in note 30.
54. This technique results in a series superior to those of Tooke (in Arnon 1991, p. 159) and Hawtrey (1918, p. 64).
55. For 1791–1800 and 1801–1815, by construction, $\sum_{i=1}^{n} F_i^B = \Delta_0^n \mathrm{SP}$.
56. For 1819–1827, QB data are beginning of subsequent year.
57. Thus, the Capie-Webber series is corrected for, inconsistently, excluding Banking Department coin but including its notes.

References

Arnon, A. (1991), *Thomas Tooke*. Ann Arbor: University of Michigan Press.

Balke, N. S., and Gordon, R. J. (1989), "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy* 97, 38–92.

Bank of England, *Quarterly Bulletin* (June 1967, Appendix).

Basu, S., and Taylor, A. M. (1999), "Business Cycles in International Historical Perspective." *Journal of Economic Perspectives* 13, 45–68.

Bayley, R. A. (1881), *National Loans of the United States*. Washington, DC: Government Printing Office.

Berry, T. S. (1943), *Western Prices before 1861*. Cambridge, MA: Harvard Univ. Press.

______ (1984), *Early California Gold, Prices, Trade*. Richmond, VA: Bostwick.

______ (1988), *Production and Population Since 1789*. Richmond, VA: Bostwick.

Board of Governors of the Federal Reserve System (1943), *Banking and Monetary Statistics, 1914–1941*. Washington, DC.

Bordo, M. D. (1975), "The Income Effects of the Sources of Monetary Change: A Historical Approach." *Economic Inquiry* 13, 505–525.

Bordo, M. D., and Kydland, F. E. (1995), "The Gold Standard as a Rule: An Essay in Exploration." *Explorations in Economic History* 32, 423–464.
Bordo, M. D., and Rockoff, H. (1996), "The Gold Standard as a 'Good Housekeeping Seal of Approval.'" *Journal of Economic History* 56, 389–428.
Bordo, M. D., and Schwartz, A. J. (1999), "Monetary Policy Regimes and Economic Performance: The Historical Record." In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, vol. 1A. Amsterdam: Elsevier.
Brezis, E. S. (1995), "Foreign Capital Flows in the Century of Britain's Industrial Revolution: New Estimates, Controlled Conjectures." *Economic History Review* 48, 46–67.
Bureau of the Census (1949), *Historical Statistics of the United States, 1789–1945*. Washington, DC: Government Printing Office.
______ (1975), *Historical Statistics of the United States, Colonial Times to 1970*. Washington, DC: Government Printing Office.
Cagan, P. (1965), *Determinants and Effects of Changes in the Stock of Money,* 1875–1960. New York: Columbia Univ. Press.
Capie, F., and Webber, A. (1985), *A Monetary History of the United Kingdom, 1870–1982*. London: Allen & Unwin.
Carothers, N. (1930), *Fractional Money*. New York: Wiley.
Department of Commerce (1986), *The National Income and Product Accounts of the United States, 1929-82: Statistical Tables*. Washington, DC: Government Printing Office.
Director of the Mint (various years), *Annual Report*.
Elliot, J. (1845), *The Funding System of the United States and of Great Britain*. Washington, DC: Blair & Rives.
Engerman, S. L. (1970), "A Note on the Economic Consequences of the Second Bank of the United States." *Journal of Political Economy* 73, 725–728.
Engerman, S. L., and Gallman, R. E. (1982), "U.S. Economic Growth, 1783–1860." In P. Uselding (Ed.), *Research in Economic History*, Vol. 8. Greenwich, CT: JAI.
Feinstein, C. H. (1988), "National Statistics, 1760–1920." In C. H. Feinstein and S. Pollard (Eds.), *Studies in Capital Formation in the United Kingdom, 1750–1920*. Oxford, UK: Clarendon.
Fenstermaker, J. V. (1965), *The Development of American Commercial Banking: 1782–1837*. Kent, OH: Kent State Univ., Bureau of Economic and Business Research.
Fenstermaker, J. V., and Filer, J. E. (1986), "Impact of the First and Second Banks of the United States and the Suffolk System on New England Bank Money: 1791–1837." *Journal of Money, Credit and Banking* 18, 28–40.
Friedman, M., and Schwartz, A. J. (1963), *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton Univ. Press.

Friedman, M., and Schwartz, A. J. (1970), *Monetary Statistics of the United States*. New York: National Bureau of Economic Research.
Gallatin, A. (1831), *Considerations on the Currency and Banking System of the United States*. Philadelphia: Carey & Lea.
Garber, P. M. (1986), "Interwar Movements of Dollars to Europe and the U.S. Currency Supply." *Journal of International Money and Finance* 5, S135–S156.
Gherity, J. A. (1993), "Interest-Bearing Currency: Evidence from the Civil War Experience." *Journal of Money, Credit and Banking* 25, 125–131.
Hammond, B. (1947), "Jackson, Biddle, and the Bank of the United States." *Journal of Economic History* 7, 1–23.
______ (1957), *Banks and Politics in America*. Princeton, NJ: Princeton Univ. Press.
Hawtrey, R. G. (1918), "The Bank Restriction of 1797." *Economic Journal* 28, 52–65.
Herfindahl, O. C. (1966), "Development of the Major Metal Mining Industries in the United States from 1839 to 1909." In *Output, Employment, and Productivity in the United States after 1800*. New York: National Bureau of Economic Research.
Highfield, R. A., O'Hara, M., and Smith, B. (1996), "Do Open Market Operations Matter? Theory and Evidence from the Second Bank of the United States." *Journal of Economic Dynamics and Control* 20, 479–519.
Holdsworth, J. T., and Dewey, D. R. (1910), *The First and Second Banks of the United States*. (Senate Document No. 571, 61st Congress, 2nd session). Washington, DC: U.S. National Monetary Commission.
Hoover, E. D., and Taylor, G. R. (1959), "Statements and Source Materials on Prices." In *Employment, Growth, and Price Levels*. Washington, DC: Hearings before the Joint Economic Committee, 86th Congress, 1st session.
Imlah, A. H. (1958), *Economic Elements in the Pax Britannica*. Cambridge, MA: Harvard Univ. Press.
Joines, D. H. (1985), "Deficits and Money Growth in the United States: 1872-1983." *Journal of Monetary Economics* 16, 329–351.
Kemp, D. S. (1975a), "A Monetary View of the Balance of Payments." Federal Reserve Bank of St. Louis *Review* 57, 14–22.
______ (1975b), "Balance-of-Payments Concepts: What Do They Really Mean?" Federal Reserve Bank of St. Louis *Review* 57, 14–23.
______ (1976), "U.S. International Trade and Financial Developments in 1976." Federal Reserve Bank of St. Louis *Review* 58, 8–14.
Kindahl, J. K. (1961), "Economic Factors in Specie Resumption: The United States, 1865–79." *Journal of Political Economy* 69, 30–48.
Laurent, R. D. (1974), "Currency in Circulation and the Real Value of Notes." *Journal of Money, Credit and Banking* 6, 213–226.
Leacy, F. H. (1983), *Historical Statistics of Canada*. Ottawa: Statistics Canada.

Martin, D. A. (1973), "1853: The End of Bimetallism in the United States." *Journal of Economic History* 33, 825–844.
______ (1976), "U.S. Gold Production Prior to the California Gold Rush." *Explorations in Economic History* 4, 437–449.
Meltzer, A. H. (1986), "Some Evidence on the Comparative Uncertainty Experienced under Different Monetary Regimes." In C. D. Campbell and W. R. Dougan (Eds.), *Alternative Monetary Regimes*. Baltimore, MD: Johns Hopkins Univ. Press.
Metcalf, C., Redish, A., and Shearer, R. (1996), *New Estimates of the Canadian Money Stock: 1871–1967* (Discussion Paper No. 96–17). Vancouver: Univ. of British Columbia, Department of Economics.
Miron, J. A. (1989), "The Founding of the Fed and the Destabilization of the Post-1914 US Economy." In M. De Cecco and A. Giovannini (Eds.), *A European Central Bank?* Cambridge, UK: Cambridge Univ. Press.
Mitchell, B. R. (1988), *British Historical Statistics*. Cambridge, UK: Cambridge Univ. Press.
Mitchell, W. C. (1908), *Gold, Prices, and Wages under the Greenback Standard*. Berkeley: Univ. of California Press.
North, D. C. (1960), "The United States Balance of Payments, 1790–1860." In W. N. Parker (Ed.), *Trends in the American Economy in the Nineteenth Century*. Princeton, NJ: Princeton Univ. Press.
Officer, L. H. (1996), *Between the Dollar-Sterling Gold Points*. Cambridge, UK: Cambridge Univ. Press.
Perkins, E. J. (1994), *American Public Finance and Financial Services, 1700–1815*. Columbus: Ohio State Univ. Press.
Redlich, F. (1968), *The Molding of American Banking*. New York: Johnson.
Report [1819] (1968), Appendix to Second Report. In *Reports from Secret Committees of the House of Commons and the House of Lords on the Expediency of the Resumption of Cash Payments* (British Parliamentary Papers, Monetary Policy, General, vol. 2). Shannon, Ireland: Irish Univ. Press.
Report [1832] (1968), Appendix. In *Report from the Committee of Secrecy on the Bank of England Charter* (British Parliamentary Papers, Monetary Policy, General, vol. 4). Shannon, Ireland: Irish Univ. Press.
Report [1840] (1968), Appendix. In *Report from the Select Committee on Banks of Issue* (British Parliamentary Papers, Monetary Policy, General, vol. 5). Shannon, Ireland: Irish Univ. Press.
Report [1841] (1968), Appendix. In *First and Second Reports from the Select Committee of Secrecy on Banks of Issue* (British Parliamentary Papers, Monetary Policy, General, vol. 6). Shannon, Ireland: Irish Univ. Press.
Report [1848] (1968), *Appendix and Index to the Reports from the Secret Committee on Commercial Distress*, Part 2 (British Parliamentary Papers,

Monetary Policy, Commercial Distress, vol. 2). Shannon, Ireland: Irish Univ. Press.

Report [1857] (1969), *Appendix and Index to the Report from the Select Committee on Bank Acts* (British Parliamentary Papers, Monetary Policy, General, vol. 8). Shannon, Ireland: Irish Univ. Press.

Rockoff, H. (1971), "Money, Prices, and Banks in the Jacksonian Era." In R. W. Fogel and S. L. Engerman (Eds.), *The Reinterpretation of American Economic History*. New York: Harper & Row.

______ (2000), "Review Essay: Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960*." EH.Net Review.

Romer, C. D. (1989), "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908." *Journal of Political Economy* 97, 1–37.

Rutner, J. L. (1974), *Money in the Antebellum Economy: Its Composition, Relation to Income, and Its Determinants*. Unpublished Ph.D. dissertation, Univ. of Chicago.

Seaman, E. C. (1852), *Essays on the Progress of Nations*. New York: Scribner.

Secretary of the Treasury (various years), *Annual Report*.

Shetler, D. (1973), "Monetary Aggregates Prior to the Civil War: A Closer Look." *Journal of Money, Credit and Banking* 5, 1000–1006.

Shrigley, I. (1935), *The Price of Gold*. London: King.

Shultz, W. J., and Caine, M. R. (1937), *Financial Development of the United States*. New York: Prentice Hall.

Simon, M. (1960), "The United States Balance of Payments, 1861–1900." In W. N. Parker (Ed.), *Trends in the American Economy in the Nineteenth Century*. Princeton, NJ: Princeton Univ. Press.

Smith, W. B. (1953), *Economic Aspects of the Second Bank of the United States*. Cambridge, MA: Harvard Univ. Press.

Solomou, S., and Weale, M. (1991), "Balanced Estimates of UK GDP 1870–1913." *Explorations in Economic History* 28, 54–63.

Solomou, S., and Weale, M. (1996), "UK National Income, 1920-1938: The Implications of Balanced Estimates." *Economic History Review* 49, 101–115.

Stevens, E. J. (1971), "Composition of the Money Stock Prior to the Civil War." *Journal of Money, Credit and Banking* 3, 84–101.

Studenski, P., and Krooss, H. E. (1963), *Financial History of the United States*. New York: McGraw-Hill.

Sylla, R. (1982), "Monetary Innovation and Crises in American Economic History." In P. Wachtel (Ed.), *Crises in the Economic and Financial Structure*. Lexington, MA: Heath.

Taus, E. R. (1943), *Central Bank Functions of the United States Treasury, 1789–1941*. New York: Columbia Univ. Press.

Temin, P. (1969), *The Jacksonian Economy*. New York: Norton.

The Economist (various issues).
Timberlake, R. H. (1993), *Monetary Policy in the United States*. Chicago: Univ. of Chicago Press.
Treasury Department (1915), *Information Respecting United States Bonds, Paper Currency, and Coin*. Washington, DC: Government Printing Office.
Urquhart, M. C. (1993), *Gross National Product, Canada, 1870–1926*. Montreal: McGill-Queen's Univ. Press.
Wallich, H. C. (1950), *Monetary Problems of an Export Economy*. Cambridge, MA: Harvard Univ. Press.
Warburton, C. (1962), "Monetary Disturbances and Business Fluctuations in Two Centuries of American History." In L. B. Yeager (Ed.), *In Search of a Monetary Constitution*. Cambridge, MA: Harvard Univ. Press.
Warren, G. F., and Pearson, F. A. (1935), *Gold and Prices*. New York: Wiley.
Weiss, T. (1992), "U.S. Labor Force Estimates and Economic Growth, 1800–1860." In R. E. Gallman and J. J. Wallis (Eds.), *American Economic Growth and Standards of Living before the Civil War*. Chicago: Univ. of Chicago Press.
Wettereau, J. O. (1937), "New Light on the First Bank of the United States." *Pennsylvania Magazine of History and Biography* 61, 263–285.
______ (1985), *Statistical Records of the First Bank of the United States*. New York: Garland.
Weymark, D. M. (1995), "Estimating Exchange Market Pressure and the Degree of Exchange Market Intervention for Canada." *Journal of International Economics* 39, 273–295.
______ (1997a), "Measuring Exchange Market Pressure and Intervention in Two Interdependent Economies: A Two-Country Model." *Review of International Economics* 5, 72–82.
______ (1997b), "Measuring the Degree of Exchange Market Intervention in a Small Open Economy." *Journal of International Money and Finance* 16, 55-79.
______ (1998), "A General Approach to Measuring Exchange Market Pressure." *Oxford Economic Papers* 50, 106–121.
Woodward, G. T. (1995), "Interest-Bearing Currency: Evidence from the Civil War Experience." *Journal of Money, Credit and Banking* 27, 927–937.

CHAPTER 23

Afterword to Part VI

23.1 Bullionist Periods (Chapter 20)

Chapter 20 was originally published in 2008, in *The Second Palgrave*.[1] Much of the chapter is based on Officer (2000). A later work (Officer 2007a) extends the latter study in various ways. Together, the three publications compose a set of innovations, or at least contributions. What are these characteristics, and how well have they held up since that time?

First, the three clearest bullionist periods (Swedish, English, Irish) are considered together.[2] Subsequent pertinent studies are Joshua R. Hendrickson (2018, 2020) and Nils Herger (2020). Hendrickson (2018) is restricted to the English experience, while Hendrickson (2020) refers to all three periods.[3] Herger's work is confined to the Swedish period.

Second, alternate models (bullionist, anti-bullionist, and their variants or offshoots) are presented via "chains of causation." Only Herger adopts that technique of stringing theories together to compose a model.[4] Third, contemporary empirical investigations—or, more broadly, studies antedating modern time-series analysis—are included in the survey. This is done by both later authors.[5]

Fourth, the data contributions of Officer (2000), confined to the Bank Restriction Period, are accepted by Hendrickson[6] Fifth, both Hendrickson and Herger are critical of time-series techniques in Officer (2000), although Herger partly follows Officer's methodology. Sixth,

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_23

Officer's econometric findings of strong evidence in favor of the anti-bullionist model combined with mixed results for the bullionist position are reversed.[7]

Seventh, perhaps for lack of awareness, Hendrickson ignores the innovations in Officer (2007a): new data series (corrected mint parity between the pound and Flemish schilling, London-Hamburg "peace-time" gold points), pound-depreciation measures (both prospective and retrospective), and resumption analyses (traditional and via gold points).[8]

23.2 Dollar-Sterling Exchange Market (Chapter 21)

23.2.1 Reception

Hugh Rockoff (2000, p. 941) kindly states: "Lawrence H. Officer has written a number of influential papers and a book on the efficiency of exchange-rate arbitrage under the gold standard." *Between the Dollar-Sterling Gold Points* (Officer 1996) was reviewed in prime economic-history outlets and in major general economics journals. Ranald C. Michie (1997, p. 207) has a substantial criticism, but states: "Otherwise it is an excellent piece of scholarship." After expressing "minor cavils," Angela Redish (1997, p. 2085) concludes: "Economic historians will thank Officer for his work and show their appreciation by using it." Together with the other reviewers, Richard Tilly is constructively critical: "the importance of Officer's labors still remains to be seen…It will be interesting to see whether financial historians will take up the challenge and opportunity that this new book offers."

I was flattered by the assessments of Marc Flandreau and Alan M. Taylor. "The book is a beautiful example of the best kind of analysis of market micro-structure, and deserves to be on the shelves of any serious student of the history of the dollar-sterling exchange rate" (Flandreau 1998, p. 1223). "The work leaves open some interesting doors for more sophisticated econometric analysis that could engage future scholars, but in many other respects this is the final word" (Taylor 1998). Also, Anna J. Schwartz kindly writes in the back cover of Officer (1996): "Students of Dollar-Sterling exchange rates will find this book to be a superb US-British financial history of the period 1791–1931 as well as an original statistical and analytic study of how the gold standard worked."

Culmination chapters of Officer (1996) constitute Chapter 21 of the present book. To understand the selection, a synopsis of the earlier part of the book is useful. Contributions of the book consist of data series (parity, exchange rate, gold points), exchange-market integration (exchange-rate, internal, external), and efficiency (market, regime)—the last represented in Chapter 21.

23.2.2 Data Series

23.2.2.1 Parity

Dollar-sterling par value ("parity"), denoted as **M** and expressed as dollars per pound, is treated in Officer (1996, ch. 5; 2006a). For 1791–1934 (ending January 30), the concept is "true mint parity," which is "based on the gold content of the sovereign and eagle after 1834, and for the period before 1834 on the silver content of the dollar and the world (Hamburg) gold-silver price" (Redish 1997, p. 2084). "We find that from 1837, until 1931, after its initial wavering, the dollar-pound parity rate settled at the famous 4.8665635 point for well nigh a century" (Taylor 1998). For 1939 (beginning September 5)–1978 (ending March 31), the U.K. official exchange rate (midpoint of the Bank of England fixed buying and selling rate) is taken, followed by the par value of the pound under the International Monetary Fund. A continuous series 1791–1934 is shown in Officer (1996, Table 5.2), and the extended annual equivalent series 1791–1978 (except for 1935–1938) listed in Officer (2006a, series Ee612).

23.2.2.2 Exchange Rate

The dollar-pound exchange-rate data are presented in Officer (1996, ch. 6; 2006b). Taylor (1998) provides a good summary:

> Officer ... present[s] data on the dollar-pound exchange rate for the entire period at quarterly frequency. In addition, monthly series are constructed for some periods: 1890-1906; 1925-1931; and, for a Bretton-Woods era comparison, 1950-66. Pre-1879 great care is taken ...to adjust the bills of exchange to a uniform zero ("sight") maturity. This ensures temporal consistency with the later cable rates; it also reflects the ultimate dominance of the sight bill as an instrument in the 1879-1914 heyday of the gold standard...Care is also taken to find a mid-point of the buy-sell rates...and further care to correct the exchange rate for devaluations of paper during

> paper standard periods. This level of care exceeds previous studies, and survives testing for the consistency and homogeneity of the series. This is probably the best quality data for the dollar-sterling exchange rate we now have for the entire period; it will be an essential series for future scholars.

Also, Michie (1997, p. 206) writes: "…the price of sterling bills of exchange in the United States. This revises the work of others, through detailed attention to the type of bill, frequency of observation, place of quotation, etc. Throughout there is a striving for accuracy that has to be admired."

The periods for quarterly data are shown in the first column of Table 23.1. The exchange-rate series, described as the "sight-equivalent exchange rate," is listed in annual form in three complementary series in Officer (2006b):

Table 23.1 Exchange-rate statistics (percentage points of parity)

Period	*Mean*		*Standard deviation*	
	Algebraic values (mean **R***)*	*Absolute values (mean* |**R**|*)*	*About mean (stdev* **R***)*	*About zero [sqrt {$\sum$***R**^{2}*/(N − 1)}]*
1791–1800	−2.70	4.55	5.06	5.75
1801–1810	3.46	4.17	3.48	4.93
1811–1820	0.97	4.57	6.19	6.27
1821–1830	1.23	2.01	2.06	2.40
1831–1840	−0.72	1.47	1.87	2.01
1841–1850	−0.73	1.11	1.26	1.46
1851–1860	0.42	0.65	0.68	0.80
1861–1870	0.32	0.87	1.20	1.25
1871–1880	−0.16	0.37	0.44	0.47
1881–1890	−0.19	0.33	0.36	0.41
1891–1900	0.02	0.25	0.30	0.51
1901–1910	−0.03	0.14	0.19	0.19
1911–1914[b]	−0.04	0.12	0.15	0.15
1919[c]–1925[b]	−0.12	0.24	0.27	0.29
1925[d]–1931[b]	−0.14	0.22	0.20	0.25
1950–1966	0.02	0.26	0.32	0.32

[a]Percent sterling premium over parity, with a negative sign denoting a sterling discount
[b]Second quarter
[c]Fourth quarter
[d]Third quarter
Source Officer (1996, Table 7.1)
R = exchange rate, percentage deviation from parity, quarterly data

Ee618 without correction for paper-currency depreciation, dollars per pound, denoted as **RS**.

Ee619 without correction for paper-currency depreciation, percentage deviation from parity, denoted **S** and computed **S** = 100 *[(**RS** – **M**)/**M**]. The monthly **S** series is plotted as "Exchange rate" in Figs. 21.1–21.3 in Chapter 21.

Ee620 corrected for paper-currency depreciation, percentage deviation from parity, denoted

R and computed **R** = **S** − **A**, where **A** is the adjustment term.[9] **R** is a series with the counterfactual basis that both countries remained on a specie standard throughout 1790–1931 (ending second quarter), excluding 3Q 1914 to 3Q 1919.

23.2.2.3 Gold Points

Gold points (generally, "specie points") differ for gold-point arbitrage (GPA) and gold-effected transfer of funds (GTF). While both involve purchasing specie in one country and transporting it to be sold in the other country, only GPA also includes a direct foreign-exchange transaction in the other direction. GTF is less than GPA, in absolute value. Gold-point estimates, with the United States as the domestic country, are generated in Officer (1996, ch. 9)—and Taylor (1998) again has a nice synopsis:

> To construct gold points requires information on costs of freight, insurance, brassage, knowledge of any gold devices used by the monetary authorities, and interest costs due to the time delay of shipment across the Atlantic Ocean. All of these are put together with the same thoroughness as the exchange rate data. The care taken places these estimates on a far firmer footing than earlier estimates which had typically cut corners...Essentially Officer proceeds with a laborious first-principles approach: each and every arbitrage cost component is individually estimated, then summed up, at each point in time. This consumes 62 pages; it is hard to imagine any improvement on these series for gold import and export points in this market, and this is the model for similar work on any other market.

Similarly, Flandreau (1998, p. 1223) writes: "the author…studies with extreme care those transformations in shipping, coining, and information technology which were liable to modify the spread of the gold points…this book will become a reference in the analysis of UK-US gold points."

Gold points may be expressed alternatively as follows.

GX = gold export point, dollars per pound
CX = cost of gold-export arbitrage or gold-effected outward transfer of funds, percent
= gold export point, percent of parity
GM = gold import point, dollars per pound
CM = cost of gold-import arbitrage or gold-effected inward transfer of funds, percent
= *minus* gold import point, percent of parity
where $\mathbf{GX} = \mathbf{M} \cdot (1 + \mathbf{CX}/100)$
$\mathbf{GM} = \mathbf{M} \cdot (1 - \mathbf{CM}/100)$

The GPA gold points estimated for the periods of the study are in Table 23.2.[10] For 1950–1966, Bank of England dollar-buying (dollar-selling) point serves as the U.S. gold export (gold import) point.[11]

23.2.3 Exchange-Market Integration

Exchange-market integration means the extent of perfection of the American foreign-exchange market. Statistics of the exchange-rate **R** offer a set of measures, shown in Table 23.1 and discussed in Officer (1996, ch. 7).[12] By any measure, there was a phenomenal improvement in integration over time.

The gold-point spread (more generally, specie-point spread, in percentage terms [percent of parity]), is the difference between the gold points: Because gold points are asymmetrical, it is useful to redefine **R**, **CX**, **CM**, **GX**, and **GM** as the percentage deviation from the midpoint of the gold-point spread. Consider the following notation:

RM = spread midpoint, dollars per pound
GS = gold-point spread, percentage points of parity
SM = spread midpoint, percentage deviation from parity
R* = **R** re-expressed as deviation from spread midpoint, percentage points of parity

Table 23.2 Gold points and exchange-market integration, Gold-point arbitrage (percentage points of parity)

Period	*Gold points*		*Integration*		
	Export *CX*	*Import* *– CM*	*Exchange-rate* *Mean* $\|R^*\|$	*External* *EI*	*Internal* *II*
1791–1800	6.1583	−8.1002	4.1854	7.1292	0.6208
1821–1830	3.6998	−4.2148	2.1443	3.9573	0.1656
1831–1840	2.3170	−4.9061	1.4438	3.6115	−0.3619
1841–1850	1.7476	−3.2960	1.0414	2.5218	−0.2195
1851–1860	1.3306	−1.8631	0.8634	1.5968	0.0650
1861–1870	1.4830	−1.4962	0.8703	1.4896	0.1255
1871–1880	1.1414	−1.0657	0.3791	1.1036	−0.1727
1881–1890	0.6585	−0.7141	0.3207	0.6863	−0.0224
1891–1900	0.6550	−0.6274	0.2470	0.6412	−0.0736
1901–1910	0.4993	−0.5999	0.1466	0.5496	−0.1282
1911–1914[a]	0.5025	−0.5915	0.1167	0.5470	−0.1568
1925[b]–1931[a]	0.6287	−0.4466	0.2764	0.5376	0.0076
1950–1966	0.6371	−0.6371	0.2564	0.6371	−0.0622

[a]Second quarter
[b]Third quarter
Source Officer (1996, Tables 9.20, 11.1)
CX = cost of gold-export arbitrage
CM = cost of gold-import arbitrage
R* = exchange rate re-expressed as deviation from spread midpoint
= **R** − (**CX** − **CM**)/2
EI = half gold-point spread
= (**CX** + **CM**)/2
II = **mean** |**R***| − **EI**/2

CX* = gold export point (**CX**) re-expressed as deviation from spread midpoint, percentage points of parity
CM* = *minus* gold import point (−**CM**) re-expressed as deviation from spread midpoint, percentage points of parity

$$\mathbf{SM} = 100 \cdot [(\mathbf{RM} - \mathbf{M})/\mathbf{M}] = (\mathbf{CX} - \mathbf{CM})/2$$
$$\mathbf{GS} = \mathbf{CX} + \mathbf{CM}$$
$$\mathbf{R}^* = \mathbf{R} - \mathbf{SM}$$
$$\mathbf{CX}^* = \mathbf{CX} - \mathbf{SM} = (\mathbf{CX} + \mathbf{CM})/2 = \mathbf{GS}/2 = -\mathbf{CM}^*$$

External integration is measured as half the gold-point spread (in percentage points of parity): **EI** = **GS**/2 = **CX***. Shown in Table 23.2, external integration is the amount of integration yet to be achieved, and this integration improves greatly over time.[13] Michie (1997, p. 206) describes external integration as "the degree of integration that existed in the exchange market…across the Atlantic." Taylor (1998) writes:

> The decline of gold point spreads mirrors that of the decline of exchange rate volatility, as expected. After 1880, this spread was at an all-time low level (even looking forward to 1925-1931 and Bretton Woods) of just above 1.0% for gold arbitrage. (Compare with around 5% in 1780, falling to about 2% in the 1840s). Officer sees this as improved "external" integration (external to the gold points) over time.

Internal integration is the amount of exchange-rate variation, measured as mean $|\mathbf{R^*}|$, given external integration. Internal integration pertains to "the internal US market" (Michie 1997, p. 206). The "expected" or "full" or "normal" value of internal integration is taken to be the midpoint of either gold point, that is, half external integration. Thus the level of internal integration, the amount of internal integration yet to be achieved, is $\mathbf{II}$ = mean $|\mathbf{R^*}| - \mathbf{EI}/2$. Internal integration over time is presented in Table 23.2 and discussed in Officer (1996, ch. 11). A negative value means that internal integration is "overfull," beyond the norm. Again quoting Taylor (1998):

> For the criterion of "internal integration" as Officer terms it, the focus is on whether "on average" the deviation of the exchange rate from parity is less than half the gold point spread, looking at absolute deviations. Again, by this measure, integration rapidly increases prior to the 1870s, then holds steady. A big jump is seen in the 1820s.

As apparent in the table and noted by Flandreau (1998, p. 1224), "external integration (measured as the reduction in the gold-point spread) dominated internal integration (measured as the reduction of exchange rate movements within the spread)." Improvements in Atlantic transportation and communication trumped the roles of banks (Second Bank of the United States, House of Brown, other New York private banks, incorporated banks).

23.2.4 Market and Regime Efficiency

The stage is set for market efficiency and regime efficiency (Chapter 21). Excluded from this chapter but presented in Officer (1996, chs. 13–15) are tests of efficiency of interest arbitrage (covered and uncovered) and forward speculation, which are confined to the 1925–1931 period. Market efficiency involves the behavior of private parties in response to profit opportunities. Regime efficiency refers to maintenance of the gold standard with the existing parity.[14] Michie (1997, p. 207) describes the conclusions well:

> Apart from imparting confidence to the exchange market through a commitment to buy gold at a fixed price, governments and central banks had little role to play before 1931. In the early years it was willingness to do nothing to hinder the movement of gold that was all important, not intervention. Similarly, even in the 1925-31 period it was the operations of the brokers and dealers that made the system work, not the activities of governments or central bankers.

23.2.5 Criticisms

Taylor (1998) criticizes the study for neglecting time-series properties of the exchange rate within the spread, while Flandreau (1998, pp. 224–225) finds fault in the explicit rejection of target-zone theory (Officer 1996, p. 285, note 1).[15]

Later works apply sophisticated econometrics—and time-series analysis in particular—to the issues that I explored. Pablo T. Spiller and Robert O. Wood (1988) and Elena Goldman (2000) have results supportive of my own; but Eugene Canjels et al. [CPT] (2004), using unique arbitrage-cost modeling, suggest that the gold-point spread in Officer (1996) is too wide.[16] It should be noted that these studies pertain only to time periods within 1879–1913. While the work of CPT is impressive, they do not identify the arbitrage-cost components that are responsible for Officer's gold-point overestimation. Nor do they address Goldman's contrary finding.[17] It appears that understanding the width of the gold-point spread during 1879–1913 might benefit from further research.

Another limitation of Officer (1996) is the absence of application to the broader economic history. "What is required is a chapter in which Officer draws conclusions from his evidence and analysis, for his research has important implications for the functioning of both international monetary systems and the world economy" (Michie 1997, p. 207). "He [Officer] does not compare application of his series to economic history with that of other series in a systematic way. Missing from his study is a discussion of the historiography of how exchange markets affected the U.S. and U.K. economies in the period studied" (Richard Tilly 1998, p. 917).

23.3 U.S. Specie Standard (Chapter 22)

Chapter 22 is the only part of this book that appeared in the previous compilation (Officer, 2007b, ch. 10). As with most authors, there is one work self-considered as warranting more attention by other scholars—and this chapter is it! Richard Sylla (2007) generously comments:

> Chapter 11, "The U.S. Specie Standard, 1792–1932: Some Monetarist Arithmetic," is one that intrigued me when it first appeared in 2002, and it still does. Among other things, careful data work, a mark of all of Officer's scholarship, produces "a monetary base series that is consistent, complete in coverage, and continuous over a long period of time" [Officer 2007b, p. 185]...It is safe to say that future work in this area will have to build on, or at least contend with, Officer's data and insights. Officer himself uses the data to study eight different regimes during the 140 years covered in the study, and concludes that the classical gold standard regime (1879-1913) was superior to the others in most respects.

Sylla takes up the issue of whether the First and Second Banks of the United States were central banks:

> One intriguing argument of the chapter is that the two Banks of the United States (BUS) in early U.S. history were indeed central banks; Officer points to substantial evidence that BUS note and deposit liabilities were held as reserves by state and other banks. This is in contrast with analyses by Temin (1969) and others, which view the monetary base as specie (gold and silver) and the BUSs as very large banks but in other respects just like all the other banks in the system. Whether the two BUSs were central banks adding to the monetary base or ordinary banks operating on a specie base obviously bears on how one might model the U.S. money supply and its proximate determinants.

Consistent with my view, Michael D. Bordo (2012, p. 598) declares: "The Second Bank of the United States under Nicholas Biddle in the decade before the Bank War had developed into a first rate central bank."

In contrast, Jane Ellen Knodell (2013, 2017) argues that the Second Bank—and, by inference, presumably also the First Bank—was a "central" bank only in the context of its historical environment and not in the modern sense of "central banking." In itself that finding is not inconsistent with Chapter 22. However, contrary to the positions of Bordo and myself, she finds that "Second Bank notes…were not generally regarded as high-powered money, fully equivalent to specie, by the state banks" (Knodell 2017, pp. 99–100).

Thus Knodell (2017, pp. 3, 100, 177) states:

> Officer (2002) drew the closest analogy of all between the Second Bank and the Federal Reserve, going so far as to include the notes and deposits issued by the Second Bank in the monetary base in one version of his model of long-run monetary growth...just as the notes and deposits of the Federal Reserve make up the monetary base in contemporary arrangements. In so doing, Second Bank money is treated as a perfect specie-substitute...I define the ultimate reserve asset (sometimes called the monetary base or outside money) as specie (coined silver and gold). I do not include Second Bank liabilities in the monetary base, as Officer (2002) and Rutner (1974) did.

Knodell's research is meticulous, but there are a number of issues unresolved in this controversy. First, Knodell makes no reference to the work of Joseph van Fenstermaker (1965), an integral component of my argument. Second, it is a matter of interpretation whether the Second Bank affected the monetary base definitionally via its liabilities as well as its gold holdings (my position) or whether (beyond its gold holdings) it affected the base only by its actions (Knodell's position). I acknowledge that Knodell's thesis makes sense, but she has not destroyed my alternative analysis (Sect. 22.1.2), let alone addressed all the points in that treatment.

Third, *if* one accepts the traditional dichotomy between high-powered and other money, then it is prima facie logical to include Second Bank notes (and arguably also the Bank's non-Treasury deposit liabilities) along with specie in the monetary base. The substitution with gold may not be perfect, but is arguably much closer than the substitution with state banknotes.

Knodell also objects to my finding that "the gold-standard period (1879–1913) outperformed the Second Bank period in terms of price stability, per capita income growth, and exchange market pressure," on grounds that "there were several banking crises during the US gold standard period, a criteria not considered in Officer's paper" and that "the Second Bank period is defined, poorly, as 1817–1838, which includes two years of monetary instability after the Bank's federal charter ended, and it continued business as the Bank of the United States of Pennsylvania" (Knodell, 2017, p. 131, note 1). Her statement is irrefutable, but the implications of the two grounds appear opposite in direction.[18]

Notes

1. Durlauf and Blume (2008).
2. Arguably, there was also a French bullionist period—during the 1790s and incorporating the famous "assignats" as well as other inconvertible currencies. See Officer (1982, pp. 42–43). In fact, that study considers the Swedish, French, English, Irish bullionist periods sequentially, albeit from the standpoint of purchasing power parity (Officer 1982, chs. 4–5).
3. It is strange that Hendrickson (2020), though largely a review of bullionist empirical evidence, makes no mention of my comparable survey (Chapter 20—originally published in 2008). Yet Hendrickson (2018, p. 211) does reference that survey.
4. His description (terse compared to mine) is "denoting a theoretical causality between economic variables by → "(Herger 2020, p. 921).
5. Hendrickson's earlier work defers to my survey. "This literature [antedating modern time-series analysis] is not summarized here both because much of this literature consists of subjective interpretations of charts and tabular data and because it is summarized adequately elsewhere (Officer 2008)" (Hendrickson, 2018, p. 211).
6. The data series are listed in Officer (2007a, pp. 266–267).
7. "There is little support for the Anti-Bullionist position. The evidence supports the Bullionist position" (Hendrickson 2018, p. 236). "In particular, the results… suggest that the increase in paper money…did cause a significant upsurge in inflation, and a depreciation of the exchange rate. Conversely, it is less clear whether non-monetary factors, such as balance-of-payments deficits, played a major role in these developments" (Herger 2020, p. 934).
8. Pamfili M. Antipa (2016) provides another analysis of resumption.
9. For paper-pound periods (1797–1821, 1919–1925), **A** is the percent currency premium over specie [typically negative, because a discount].

For paper-dollar periods (1814–1817, 1837–1842, 1857, 1862–1878), **A** is the percent specie premium over currency [typically positive].

10. For GTF (corresponding to GPA) gold points, see Officer (1996, Table 9.20).
11. Monthly GPA gold points are graphed in Figs. 16.1–16.3.
12. Table 23.1 is reproduced in Eugene Canjels et al. (2004, p. 869). The symbolic headings are introduced here.
13. Discussion of external integration is in Officer (1996, ch. 10). External integration is half the spread, for convenience in defining internal integration.
14. Nuno Palma and Liuyan Zhao (2021, pp. 893–897) apply the regime-efficiency model to the Chinese silver standard.
15. C. Paul Hallwood et al. (1996) had already investigated the extent to which gold standards were "well-behaved target zones."
16. Palma and Zhao (2021, p. 897) find favor in the CPT results.
17. "Officer's conclusion of a 'remarkably efficient' gold standard between 1890 and 1906…is confirmed by our estimate of unit root incorporating double truncation" (Goldman, 2000, p. 258).
18. Bordo (2012, p. 606), in counterfactual analysis, demonstrates that "had the Second Bank of the United States not been destroyed in 1836 that the US could have had a better history with respect to financial stability, price stability and overall macro-performance." This finding is in line with Knodell's sentiment.

References

Antipa, Pamfili M. (2016). "How Fiscal Policy Affects Prices: Britain's First Experience with Paper Money." *Journal of Economic History* 76 (December), pp. 1044–1077.

Bordo, Michael D. (2012). "Could the United States Have Had a Better Central Bank? An Historical Counterfactual Speculation." *Journal of Macroeconomics* 34 (September), pp. 597–607.

Canjels, Eugene, Gauri Prakash-Canjels, and Alan M. Taylor (2004). "Measuring Market Integration: Foreign Exchange Arbitrage and the Gold Standard, 1879–1913." *Review of Economics and Statistics* 86 (November), pp. 868–882.

Durlauf, Steven N., and Lawrence E. Blume, eds. (2008). *The New Palgrave Dictionary of Economics*, 2nd ed. Houndmills, Basingstoke, Hampshire and New York, eight vols.

Fenstermaker, Joseph Van (1965). *The Development of American Commercial Banking: 1782–1837*. Kent, OH: Kent State University, Bureau of Economic and Business Research.

Flandreau, Marc (1998). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. *Economic Journal* 108 (July), pp. 1223–1225.

Goldman, Elena (2000). "Testing Efficient Market Hypothesis for the Dollar-Sterling Gold Standard Exchange Rate 1890–1906: MLE with Double Truncation." *Economics Letters* 69 (December), pp. 253–259.

Hallwood, C. Paul, Ronald MacDonald, and Ian W. Marsh (1996). "Credibility and Fundamentals: Were the Classical and Interwar Gold Standards Well-Behaved Target Zones?" In Tamim Bayoumi, Barry Eichengreen, and Mark P. Taylor, eds., *Modern Perspectives on the Gold Standard*, pp. 129–161. Cambridge: Cambridge University Press.

Herger, Nils (2020). "An Empirical Assessment of the Swedish Bullionist Controversy." *Scandinavian Journal of Economics* 122 (July), pp. 911–936.

Hendrickson, Joshua R. (2018). "The Bullionist Controversy: Theory and New Evidence." *Journal of Money, Credit and Banking* 50 (February), pp. 203–241.

______ (2020). "Bullionism." In Stefano Battilossi, Youssef Cassis and Kazuhiko Yago, eds. *Handbook of the History of Money and Currency*, pp. 983–993. Singapore: Springer Nature.

Knodell, Jane Ellen (2013). "The Nation-Building Purposes of Early US Central Banks." *Review of Keynesian Economics* 1 (Autumn), pp. 288–299.

______ (2017). *The Second Bank of the United States*. London: Routledge.

Michie, Ranald C. (1997). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. *Economic History Review* 50 (February), pp. 206–207.

Officer, Lawrence H. (1982). *Purchasing Power Parity and Exchange Rates*. Greenwich, CT: JAI.

______ (1989). "The Remarkable Efficiency of the Dollar-Sterling Gold Standard, 1890–1906." *Journal of Economic History* 49 (March), pp. 1–41.

______ (1996). *Between the Dollar-Sterling Gold Points*. Cambridge: Cambridge University Press.

______ (2000). "The Bullionist Controversy: A Time-Series Analysis." *International Journal of Finance and Economics* 5 (July), pp. 197–209.

______ (2002). "The U.S. Specie Standard, 1792–1932: Some Monetarist Arithmetic." *Explorations in Economic History* 39 (April), pp. 113–153.

______ (2006a). "Dollar-sterling parity: 1789–1978." In Susan B. Carter and others, *Historical Statistics of the United States, Millennial Edition*, pp. 5–561 to 5–563. New York: Cambridge University Press.

______ (2006b). "Dollar-sterling exchange rates: 1791–1914." In Susan B. Carter and others, *Historical Statistics of the United States, Millennial Edition*, pp. 5–563 to 5–566. New York: Cambridge University Press.

______ (2007a). "Bank Restriction Period, 1797–1821." In Officer (2007b), pp. 253–267.

______ (2007b). *Pricing Theory, Financing of International Organizations and Monetary History*. London: Routledge.

______ (2008). "Bullionist Controversies (Empirical Evidence)." In Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, second edition, vol. 1, pp. 597–602. Houndmills and New York: Palgrave Macmillan.

Palma, Nuno, and Liuyan Zhao (2021). The Efficiency of the Chinese Silver Standard, 1920–1933." *Journal of Economic History* 81 (September), pp. 872–908.

Redish, Angela (1997). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. *Journal of Economic Literature* 35 (December), pp. 2084–2085.

Rockoff, Hugh (2000). "Banking and Finance, 1789–1914: Bibliographical Essay." In Stanley L. Engerman and Robert E. Gallman, eds., *The Cambridge Economic History of the United States*, vol. 2, *The Long Nineteenth Century*, pp. 933–945. Cambridge: Cambridge University Press.

Rutner, J. L. (1974). *Money in the Antebellum Economy: Its Composition, Relation to Income, and Its Determinants*. Unpublished Ph.D. dissertation, University of Chicago.

Spiller, Pablo T., and Robert O. Wood (1988). "Arbitrage during the Dollar-Sterling Gold Standard, 1899–1908: An Econometric Approach." *Journal of Political Economy* 96 (August), pp. 882–892.

Sylla, Richard (2007). Review of Lawrence H. Officer, *Pricing Theory, Financing of International Organizations and Monetary History*. EH.net.

Tilly, Richard (1998). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. *Journal of Economic History* 58 (September), pp. 916–917.

Taylor, Alan M. (1998). Review of Lawrence H. Officer, *Between the Dollar-Sterling Gold Points*. EH.net.

PART VII

Economics in Alternative Scenarios

CHAPTER 24

Economics and Economic History in Science Fiction

Ari J. Officer and Lawrence H. Officer

24.1 What Is Science Fiction?

Science fiction—and speculative fiction more broadly—is not primarily about bug-eyed monsters or wizard fantasies or even the future. Rather, science fiction presents alternate realities, ranging from natural extensions of the present time to tales of the near future, to stories taking place in the distant future—on Earth and beyond. Counterfactual history, well established in the economic-history literature, traces its roots to the same literary genre.[1] Speculative fiction reflects history, sometimes with novel ways of viewing the past. Even far-future fiction can reimagine historical events and trends. Most important, science fiction branches off from history. Whether it extrapolates from the modern world or some point in the past, science fiction reflects human society and places people into new situations.

Such circumstances are worth studying, as they shed light on social sciences such as economics. This chapter reviews science-fiction literature and film that incorporate economic topics in a historical context. We reflect on the authors' and filmmakers' views as revealed in their fictional economic systems. This is not a complete study but rather a targeted

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8_24

sampling of seminal works, organized, in Sects. 24.2–24.7, into Technology, Labor, Assets and Capital, Trade, Government, and Alternate History. The chapter concludes with Sect. 24.8, a plea for recognition of the importance of science fiction.[2]

As publishing has become more meritocratic and accessible, authors are releasing new and exciting science fiction every day. The study of the genre in this chapter hopefully justifies further such analysis of more works, especially those from emerging authors of diverse backgrounds whose perspectives were not as readily published or promoted in the past.

Because civilization typically evolves slowly, academic study often suffers from "status quo bias." As history moves continuously, there is a tendency to restrict the scope of real-world data to established convention. In economics, even controlled experiments analyze behavior inherently skewed by the subjects' experiences. Therefore, "thought experiments" occurring outside the present reference frame can aid understanding and play an essential role in economic research.

The catalog of science fiction provides an abundance of thought experiments far removed from the present: How would economic frameworks work in other worlds? Do any inconsistencies in the conventional practice or theory become more obvious by twisting reality? How does economic history facilitate the understanding of authors' worldbuilding decisions?

Science fiction, then, comes down to setting. Any complete setting must have an underlying economic system, in part to justify character motivation. Such economies generally reflect those of history, with interesting divergences resulting from the tweaks to reality.

For many years, the "back page" of the *Journal of Political Economy* has presented passages from literature to illustrate economic principles, including excerpts from science fiction suggested by one of us.[3]

Reflecting an author's time and society, literature often responds to social and political movements. Science fiction, in particular, gives insight into cultures hostile to open criticism. It provides data that authoritarian governments might otherwise suppress. Science fiction creates a space with less bias, where readers do not feel as personally invested in the world—because it is fictional. This space allows readers to draw conclusions or realize concerns they would not necessarily see otherwise.

While fiction cannot "prove" any theory, it can help illuminate potential flaws and shortcomings. Studying speculative fiction, then, enhances both economic theory and economic practice. There are often unintended consequences of government policy, especially in a social science tied to human behavior. People do not always act as expected in theoretical models. We suggest that, in these times of increasing government control and intervention in financial and labor markets, policymakers would benefit from studying science fiction and thereby gaining both more imagination and more creativity!

A final introductory thought: Speculative fiction often predicts the future, pondering "what might be." But often more useful and interesting is to consider *what might go wrong*.

24.2 Technology

Technology often plays a crucial role in science fiction. So-called "hard science fiction" tends to create futuristic settings with new science or technology that is integral to the plot. Often the inciting incident or ultimate solution to the protagonist's problem involves the use of technology; otherwise, the story could have occurred in a non-speculative setting.

If technology is integral to a story's plot, there should also be economic consequences, either front-and-center or implied. As technology drives economic growth—from the printing press to self-driving cars, among countless historical innovations—a fictional-world's economy should reflect its unique set of technology.

24.2.1 Space Travel

Many of the advances postulated in fiction have already come to fruition; thus, the most incisive ones are those humanity has not yet fully realized. At the present time, space travel has already enhanced the terrestrial economy through satellites and various scientific breakthroughs. Extrapolating this progress—and even writing prior to the Space Race—many authors imagined distant futures in which humankind will colonize not only the solar system but also distant parts of the galaxy.

Isaac Asimov's classic *Foundation* encapsulates economic principles on multiple levels. A future society such as Asimov's vast Galactic Empire would be incomplete without a well-formulated economy. Character motivation would fall apart without economic incentive.

Asimov introduces a fictional statistical sociology called "psychohistory," which seeks to explain and predict the behavior of the mass of humanity. This study leads its developer to predict that, as with all empires historically, the Galactic Empire will fall apart: “Interstellar wars will be endless; interstellar trade will decay; population will decline” (Asimov 2004, p. 30). The novel and its sequels follow the people committed to protecting humanity’s knowledge and shortening the period of turmoil following the Galactic Empire’s collapse.

Society has never undergone a war across light-years of distance. The uncertainties associated with such a war, which could span centuries, would suppress innovation. The collapse of trade would impair the economy of every planet. Instead of specializing in the production of goods and services most suited to the population and available resources, the economy would instead produce other goods less efficiently than its trading partners. Even those industries in which a planet excels would become uneconomical without the support of foreign purchases: “If Korell [a trade partner] prospered with our trade, so did we. If Korellian factories fail without our trade; and if the prosperity of the outer worlds vanishes with commercial isolation; so will our factories fail and our prosperity vanish” (Asimov 2004, p. 243). A breakdown in trade can cause a deleterious chain reaction in an economy, particularly when that economy has become so specialized that it relies on imports for low-priced necessities. “Mutual gains from trade” has long been a fundamental theorem of economics, and it is pleasing that a prolific science-fiction writer sees not only that theorem but also the obverse.

Worldwide economic disruptions, such as the Great Depression of the 1930s, involve tremendous declines in international trade. Generally, the causation is two-way (“the international business cycle”).

International migration—as distinct from trade in commodities—can be countercyclical. Economic downturns in some countries historically led to increased migration to other countries. Episodic bulges in the movement of people from Latin America to the United States are examples. The ultimate causation is various: political corruption, lack of freedom (dictatorship, junta rule), domestic terrorism (drug cartels, gang warfare, breakdown of law and order), natural disasters (climate change, hurricanes). Much international movement of people occurs because of war—between India and Pakistan after the partition in 1947, from Ukraine to Poland and other neighboring countries after the Russian invasion in 2022.

Concentrated agrarian economies are especially susceptible to disruption. Because certain crops have been cultivated to a predominant species that produces the maximum yield, regions that rely on a single crop have failed to support themselves in the event of drought or crop blight. Lack of biodiversity can be a cause of famine. In particular, the Irish Great Famine of the late 1840s resulted in part from genetically similar potatoes succumbing to mold. To escape starvation, a massive Irish migration to other countries ensued, especially to the United States.

Today's global trade has resulted in intertwined economies, even more so than when Asimov penned *Foundation*. Disputes between two countries can result in economic harm to them both. Conflict does not have to be physical; cold wars, as well as trade wars, impair both economies. Tariffs and local subsidies generally reduce competition and harm consumers.

What happens when trade breaks down? Just as it occurs in the future in Asimov's *Foundation*, it can take place in our time on Earth. Recent events have exposed multiple issues in the global supply chain. Trade disruptions can delay the production and distribution of consumer and producer goods both.

While Asimov's psychohistory focuses on averages of large groups of people over time, other works have explored the economic implications of individual decisions. In Frederik Pohl's *Gateway*, humanity has discovered long-abandoned starships from an ancient alien civilization. But navigating the galaxy with those ships involves substantial risk: "You literally had no control, once you started out in a Heechee [alien] ship. Their courses were built into their guidance system, in a way that nobody had figured out; you could pick one course, but once picked that was it—and you didn't know where it was going to take you" (Pohl 1999, p. 26). More importantly, the explorers do not know a voyage's length. Given that the ships have to be retrofitted for human life support, there is a limit to how long someone can survive.

To further complicate the risk-reward calculus, the probability of success (discovering valuable artifacts) is extremely low: "About eighty percent of flights from Gateway came up empty. About fifteen percent didn't come back at all. So one person in twenty, on the average, comes back from a prospecting trip with something that Gateway—that [humankind] in general—can make a profit on. Most of those are lucky if they collect enough to pay their costs for getting here in the first place" (Pohl 1999, p. 40). Based on those odds, explorers are more likely to die than come out ahead financially!

Every day, people assess various decisions. Rarely are the risks and rewards as clearly defined—and extreme—as in *Gateway*. Yet there are always those desperate or ambitious enough to take the risk. People typically do not grasp (or they deem irrelevant) true probability when making decisions; if they did, many would not play the lottery! It is human nature to take chances based on the best-case scenario, rather than realistic expectation, focusing on an ultra-desirable outcome despite an extremely low probability of attainment.

Early explorers and later colonists crossed the Atlantic Ocean on sailing ships, with an uncertain future even if they survived the voyage. They took their chances, and eventually Western civilization benefited from expansion into the Americas.[4] Similarly, humanity in Pohl's universe benefited from the technologies discovered by those risking death for personal wealth.

As civilizations have risen and fallen over the ages, it is reasonable to expect that a civilization spanning the galaxy could fall apart. Furthermore, space travel without faster-than-light communication would lead to colonies effectively cut off from Earth save for multi-year communications delays. What happens to a colony if communication with Earth breaks down altogether? Roger Zelazny describes one such colony in *Lord of Light*. The crew takes advantage of the passengers, setting itself atop a Hindu caste system on the new planet, with only one crewmember believing they should act otherwise: "I felt that we should be doing something about the passengers, as well as the offspring of our many bodies, rather than letting them wander a vicious world, reverting to savagery. I felt that we of the crew should be assisting them, granting them the benefits of the technology we had preserved, rather than building ourselves an impregnable paradise" (Zelazny 2000, p. 63). As technology advances, it is likely that certain members of society will abuse it to maintain power over others.[5]

The British Parliamentary Iron Act of 1750 restricted the manufacture of iron and steel products in the American colonies, allowing colonial export only of unfinished iron—to be processed in Great Britain. By curbing the use of raw materials in the colonies, Britain would maintain economic power. However, suppressing access to technology limits productivity, and America later created opportunities for skilled immigrants to bring technologies and develop them further in the independent United States.

24.2.2 Medical Advances

Medical advances have extended lifespans and reduced illness, expanding employment. But what happens as these advances continue? Should humans be able to live a long time—if not indefinitely—in some form, consumption habits and productivity levels would inevitably change.

In Richard K. Morgan's *Altered Carbon*, human consciousness can be transferred to new bodies called "sleeves," allowing both efficient travel over large distances and prolonged human life. But the prohibitive cost of premium sleeves comparable to human bodies means most people do not benefit from the technology. Instead, more people corner fewer resources, and the economy responds in a balanced way. Circumstances can quickly change in Morgan's world: "One day you own a house, your sleeve policy's paid up, the next you're on the street looking at a single life span" (Morgan 2006, p. 123). With higher-quality goods and services, but also higher costs, both lifespan and the quality of life for the average person can drop.

As the narrator's employer says, "I wonder if you realize how much it is costing me just to keep you alive and out of storage" (Morgan 2006, p. 207). By physical existence becoming transactional, people gain direct power over others. While this is an exaggeration of current reality, it sheds some light on "medical debt" in modern society.

Since the 1960s, health-care spending in the United States has increased substantially, both absolutely and as percentage of GDP, even as the overall economy has expanded tremendously. As people live longer and more medical conditions are treatable, the increase in health-care consumption is understandable.

Because economic progress is often driven by individuals pursuing their own interests, living longer results in potential shifts in behavior: "We have a friendship that goes back centuries. Common business strategies that have sometimes taken longer than a human lifetime to bring to fruition" (Morgan 2006, p. 490). Human mortality has limited the long-term scope of any individual's horizon. Most large corporations become publicly owned across many shareholders, and even closely held private companies get split up as they are inherited.

In the film *The 6th Day*, directed by Roger Spottiswoode, humans can be physically cloned and have their memories fully restored in new bodies. Beyond the ethical and spiritual questions of nature and identity, the movie touches on economic consequences for the labor force. At the beginning of the film, an injured athlete is worth far less to a sports franchise, the manager of which decides to replace him with a clone. People, of course, have value beyond economic production; even through the economic lens, a star athlete generates value beyond playing games and has other revenue streams. But the team may not benefit from economic activity outside the athlete's contract. Unrestricted capitalism and free markets do not always value potential fully; those directly benefiting from a worker may not have the means or even the motivation to advance the worker's potential.

Developing a worker's skills to their long-term potential has not always been the revealed practice of nations or economies. The United States has drafted millions of men over the years to fight wars. Similar to the athlete's contract in *The 6th Day*, citizens have an obligation to defend their country despite conflict with their own professional interests.

While not a medical advance per se, space travel would allow an isolationism that could reduce the spread of disease. In *The Caves of Steel*, Isaac Asimov writes that the "precision with which the Spacers had bred disease out of their societies was well known. The care with which they avoided, as far as possible, contact with disease-riddled Earthmen was even better known" (Asimov 1985, p. 6). The health of colonists in space, with limited crews and mission specialization, is essential to their survival.

As international travel has become cheaper, tourism on Earth has boomed. This interconnectedness undoubtedly accelerated the COVID-19 pandemic. The 1918 influenza pandemic was exacerbated by World War I, with troops traveling and living in close quarters. While Earth is now free of naturally occurring smallpox, its spread as the Americas were colonized devastated the indigenous populations, who lacked natural immunity.

24.2.3 Robotics and Materials

Asimov imagines that colonization beyond Earth can occur only with advances in materials sciences and robotics. Specifically, "those very robots can accompany humans, smooth the difficulties of initial adjustment to a raw world, make colonization practical" (Asimov 1985, p. 97). Increased automation lowers the risk to humans and increases the chances of success. Asimov's 1953 novel came out a few years before Sputnik 1 orbited the Earth. Since then, space programs have used unmanned vessels and machines to explore the solar system.

Beyond the practical considerations, Asimov also explores the socioeconomic consequences of humanity inhabiting both the Earth and other, less-hospitable worlds: "It's the difference between us and the Spacers. We reach high and crowd close. With them, each family has a dome for itself. One family: one house. And land between each dome" (Asimov 1985, p. 5). Those seeking to colonize other worlds would surely not want to live in a dense city. Yet, on Earth, with land limited and services concentrated in cities, urbanization will likely continue.

As with the machines in Asimov's novels, Philip K. Dick's robots in *Do Androids Dream of Electric Sheep?* are developed primarily for colonization. This use creates an economic relationship such that "the manufacture of androids, in fact, has become so linked to the colonization effort that if one dropped into ruin, so would the other in time" (Dick 1982, p. 40). Given the negative perception of technology leading to job displacement, it makes sense that androids would be deployed mostly off-world.

Despite best intentions, not all technology is positive for society or the economy. The satirist Kurt Vonnegut demonstrates this phenomenon throughout his work. In *Cat's Cradle*, a scientist develops a substance called "ice-nine," which solidifies water at higher temperatures "to get Marines out of the mud" (Vonnegut 2006, p. 44). Such an innovative technology could reduce the burden on the military, allowing economic capacity to shift into other industries and uses. Unfortunately, ice-nine does not work as intended. It turns any continuous body of water into ice: "I opened my eyes—and all the sea was ice-nine" (Vonnegut 2006, p. 261). Technology can have unintended consequences for the broader civilization, as a solution to one problem can generate its own issues.

In the film *The Matrix*, directed by Lana and Lilly Wachowski, machines have trapped humanity in a simulated reality in order to use humans as an energy source. The message—contrary to Asimov's First Law of Robotics—is that advances in artificial intelligence could potentially backfire and harm society.

Before the dangers of smoking were well known, the cigarette industry expanded access to tobacco by making the substance cheaper and more convenient. Even government-backed research intended to help society has harmed it. For decades, the U.S. Department of Agriculture promoted a diet low in fat and high in carbohydrates. Consequently, food companies removed fat from foods and replaced it with sugar to maintain taste. While some businesses have benefited (such as the pharmaceutical industry, thanks to an increase in diabetes), it stands to reason that workers are less productive than if they were healthier.

24.2.4 Communication

The expansion of civilization into the stars would necessitate advances in communication to bring closer together those who are physically far apart. On Earth, too, improved communication has allowed remote work and outsourcing to maximize the use of resources across the globe. But, as with ice-nine, technologies have unintended consequences.

In the early days of personal computing, William Gibson imagined a future in which communication would *not* bring people together, depicted in *Neuromancer*: "We have sealed ourselves away behind our money, growing inward, generating a seamless universe of self" (Gibson 1984, p. 173). Before computer technology, most people needed to have human-to-human interaction to thrive, whether collaborating in person or engaging in trade with others. The digital economy, however, allows people to work remotely and avoid direct interactions.

Well before technology behemoths such as Google or Facebook existed, Gibson predicted that an information economy would result in larger, more-pervasive corporations: "Power, in Case's world, meant corporate power. The zaibatsus, the multinationals that shaped the course of human history, had transcended old barriers. Viewed as organisms, they had attained a kind of immortality. You couldn't kill a zaibatsu by assassinating a dozen key executives; there were others waiting to step up the ladder, assume the vacated position, access the vast banks of corporate memory" (Gibson 1984, p. 203). Information technology allows scalability that manufacturing companies, for instance, can never attain. It also diffuses responsibility such that the technology grows uncontrolled. Social-media algorithms, designed to maximize engagement and advertising revenue, can be used—sometimes inadvertently, sometimes deliberately—to manipulate people and spread false information.

On the other side, improved communication has legitimately brought many people closer together. Ursula K. Le Guin imagined instantaneous communication across large distances, faster than the speed of light. When her novel *The Left Hand of Darkness* was published in 1969, the Internet was in its infancy, with the first computer network having just been built between Pentagon-funded research institutions. Yet Le Guin takes the current events of her time and extrapolates them to a bright future: "The kind of trade I speak of can be highly profitable, but it consists of simple communication rather than transportation. My job here is, really, to find out if you're willing to communicate with the rest of mankind" (Le Guin 2000, p. 137). Because transportation over long distances is resource-intensive, transferring information is a more economical way to provide value through trade. Indeed, the digital economy was estimated at $11.5 trillion in 2016, far larger than the $2.7 trillion GDP of the entire world economy in the year Le Guin's novel was released.[6] Even adjusting for U.S. inflation, the digital economy has grown nearly as large as the entire economy was decades ago. Ursula K. Le Guin effectively predicted the economic benefit of the Internet before its existence!

24.3 Labor

As technology changes, so does labor. The fear of job displacement due to robotics and similar technologies has pervaded popular sentiment for decades. Thus, it is no coincidence that labor plays a role in many science-fiction books and films, as writers respond to the changing economies of their own times.

Despite short-term job losses from innovation, technology has brought new industries and occupations to the global economy. The large technology companies that William Gibson predicted now employ millions of workers around the world—in jobs that would not have otherwise existed.

24.3.1 Labor Force Participation

Le Guin not only predicted the rise of information technology, but she also foresaw the benefits of a workforce free of prejudice. *The Left Hand of Darkness* takes place on a planet populated by future humans who have evolved to have no fixed sex, occasionally becoming male or female for reproductive purposes. One of the immediate consequences of a society without distinct genders is that the workforce is not divided along those lines, which leads to increased productivity. Since 1969, the labor-force participation rate of women in the United States has increased from 43 to 56 percent.[7] Undoubtedly, the economy has grown faster as a result.

In *Player Piano*, Kurt Vonnegut describes a mechanized society in which a small elite group of engineers and managers runs the automated infrastructure, while most people have no means to earn a living: "Now, you people have engineered them out of their part in the economy, in the market place [sic]" (Vonnegut 1970, p. 92). True, disruptive industries have short-term consequences for specific parts of the labor force; but history has so far shown that new opportunities arise that improve employment in the long run. While computers displaced some jobs, they also became a pervasive part of virtually every industry.

At some point, as artificial intelligence and related technologies mature, it is possible that the world could face a reckoning in which labor has been permanently displaced. At the very least, it makes sense to consider the short-term impact to jobs of new technologies in order to help facilitate training of new skills valuable to the changing economy.

Vonnegut draws an interesting comparison: "The machines are to practically everybody what the white men were to the Indians. People are finding that, because the way the machines are changing the world, more and more of their old values don't apply any more. People have no choice but to become second-rate machines themselves, or wards of the machines" (Vonnegut 1970, p. 274). Workers have had to adapt to changes in society and the economy, both of their own volition and at the will of their government.

24.3.2 Population Growth

Changes in population levels, both growth and decline, impact the labor force—as well as the population's needs, based on age distribution. If population grows faster than the economy can harness resources, quality of life declines. Movies such as *Soylent Green*, directed by Richard Fleischer, warn of overpopulation coupled with dwindling resources complicated by climate change. Food technology has not had the same level of investment as other industries, yet food is essential for life.

Phyllis D. James's novel *Children of Men* takes on the other risk: what if the population stopped growing and reproduction became extremely rare? Amidst mass infertility, "the country sunk in apathy, no one wanting to work, services almost at a stop, crime uncontrollable, all hope and ambition lost" (James 2006, p. 152). Human behavior changes based on perceptions of the future. Without the need to provide for descendants, people do not need to work as much. Thus, there is less production from those in the workforce, as well as a lack of new workers.

Populations have historically declined during periods of disease such as the Black Death and those of hunger such as Ireland's Great Famine.

A decline in population growth, or of population itself, and declines in birth rates have harmful effects on the economy—these declines a fear of policymakers in many developed countries in the present day.

Whereas Le Guin developed a world of equality, Larry Niven and Jerry Pournelle envisioned an alien species with hyperspecialized castes tied to their physical development. *The Mote in God's Eye* introduces an alien race of multiple variants that also alternates between sexes, but with the complication that every alien "has to be made pregnant after she's been female for a while. Child, male, female, pregnancy, male, female, pregnancy, 'round and 'round. If she doesn't get pregnant in time, she dies" (Niven and Pournelle 1974, p. 302). A positive feedback loop occurs, with consumption of resources leading to even more consumption. Ultimately, the aliens are "locked into a permanent state of population explosion followed by total war" (Niven and Pournelle 1974, p. 461). Resources are limited, and thus explosive population growth cannot continue unabated.

The Niven-Pournelle worldview, as well as that of Le Guin, replaces characteristics of the present-day recognition of (and discourse about) gender dysphoria with universal gender fluidity. Adapting an aspect of humanity into a speculative element is an excellent example of an important aspect of science-fiction literature: extrapolation of a current societal trend to a logical end.

Although humans have not reached global levels of overpopulation, local conflicts historically have arisen over resources. Britain conquered India in the eighteenth century to gain access to goods such as cotton, silk, tea, and spices. The Soviet Union invaded Finland in 1939, partially over nickel mines; and Russia invaded Ukraine in 2022, in part for its resource base (including agriculture and minerals such as lithium).[8] In 1942 Hitler determined to secure the Caucasus oil fields for the German military, thus cutting off that area from the Soviet Union. And Iraq invaded Kuwait in 1990 in an attempt to control Kuwait's oil reserves.

24.3.3 Specialization

The aliens in *The Mote in God's Eye* evolved into multiple distinct subspecies with skills suited to their physiques. Among other castes, Engineers and Watchmakers excel at mechanical and electrical work; Doctors specialize in medicine and have steady hands for surgery; Farmers focus only on agriculture; and Masters make decisions for their communities and the broader race. Because the skills in each caste are tied to genetics, the aliens are born into their roles with no possibility of changing caste. Interbreeding Masters and Engineers, for example, creates sterile Mediators; the castes are destined to continue so long as each subspecies survives.

Nora K. Jemisin, on the other hand, builds a future world that periodically undergoes extreme climate change, resulting in the population forming communities called "comms." Each of these comms needs members of specific use-castes in order to survive. Similar to Doctors in *The Mote in God's Eye*, Resistants in *The Fifth Season* are thought to be immune to disease, and thus, "comms like to have them around no matter how hard the times, in case of sickness and famine and such" (Jemisin 2015, p. 27). Strongbacks, on the other hand, are the laborer caste: "All 'Strongback' means is that her female ancestors were lucky enough to join a comm but too undistinguished to earn a more secure place within it. *Strongbacks get dumped same as commless when times get hard*" (Jemisin 2015, p. 27). With trade inherited from parents, these castes are ancestral but somewhat indistinguishable, with nowhere near the genetic specialization described in *The Mote in God's Eye*. There is no social mobility, with laborers defined by their given jobs, and those jobs potentially dooming them if their comm deemed them non-essential. In this case, however, the caste system arises from culture more than from genetics.

Perception affects economic opportunity. Historically, racism and sexism limited who could more easily work in specific occupations. A similar caste system emerged in ancient India, with four main groups in a hierarchy: priests, warriors, merchants, peasants. The marginalization of certain groups continues to the present. Even without constant crisis, as in Jemisin's world, society can perpetuate a more- or less-rigid social system.

Yet, in times of crisis, identity no longer matters. During the Great Depression, cities were hit especially hard. In the United States, several million unemployed young men left cities for Civilian Conservation Corps work, just as the Strongbacks in the novel. Eventually this work ended, and the overall economy improved.

In Jemisin's world, however, natural disasters keep people from overthrowing the system. When the comms face disasters of their own, their options are limited: "These people face the economic destruction of their whole community. It's not a Season, so they can move somewhere else, try to start over. Or they can dissolve, with all the comm's families trying to find places in other communities—which should work except for those family members who are poor, or infirm, or elderly. Or those who have uncles or siblings or parents who turned out to be orogenes; nobody will take those. Or if the community they try to join has too many members of their use-caste already" (Jemisin 2015, p. 196). Who are the essential workers? Who does society need to function? The orogenes in the novel—those with supernatural power—are akin to the casteless "untouchables" in India. But they are instead marginalized for their abilities.

Bong Joon-ho's film *Snowpiercer* depicts a microcosm of society on a moving train, with the passengers segregated by socioeconomic classes similar to those in Western culture. With no social mobility and no way off the train, the working class revolts.

Forcing specialization by class has backfired many times throughout history. In the Russian Revolution of 1905, peasants and industrial workers revolted against the nobility. The peasants had been provided with small plots of land that could not generate enough food but were burdened with property taxes and maintenance costs.

24.4 Assets and Capital

24.4.1 Consumption and Inflation

War can unite a people due to both culture and survival, building up a nation's production and overall capital. This concept arises in *The Left Hand of Darkness*: "He was after something surer, the sure, quick, and lasting way to make people into a nation: war" (Le Guin 2000, p. 102).

Joe Haldeman's aptly titled *The Forever War* covers a long-term war against an alien civilization, exploring the consequences of relativistic time dilation. Because this is such a major effort, it consumes most of the economic resources back on Earth. During World War II, the United States spent over 40% of GDP on defense.[9] This was only temporary, but a long-term economic dependence on war could have intriguing consequences: "The main effect of the war on the home front was economic, unemotional—more taxes but more jobs as well. After twenty-two years, only twenty-seven returned veterans; not enough to make a decent parade. The most important fact about the war to most people was that if it ended suddenly, Earth's economy would collapse" (Haldeman 2014, p. 139). Wars can lead to increased production for defense, fostering economic expansion. After World War II, the United States produced goods for other countries to consume as they rebuilt their infrastructure after the war's destruction. Modern wars in the Middle East have benefited specific companies: those with oil interests, weapons manufacturers, and government contractors.

Contemporary soldiers, however, are not affected by relativity as those in *The Forever War*. Time dilation results in soldiers earning significant money, as time passes much faster on Earth during their deployment light-years away. Without much sales competition at their base, "Heaven's economy was governed by the continual presence of thousands of resting, recreating millionaire soldiers. A modest snack would cost a hundred bucks, a room for a night at least ten times that. Since UNEF [the armed forces] built and owned Heaven, this runaway inflation was pretty transparently a simple way of getting our accumulated pay back into the economic mainstream" (Haldeman 2014, p. 175). Inflation, in this case, results from a wealthy subsection of civilization being concentrated at an isolated base.

As in *The Forever War*, prices are not the same throughout a country. Despite being a standardized commodity, gasoline of a given grade varies in price across the United States due to local taxes and transportation costs. Also, gas stations that are more convenient and in wealthier areas can charge more than their competitors located elsewhere.

The film *In Time*, directed by Andrew Niccol, introduces a society in which the currency used is time remaining in an individual's life. Society becomes bifurcated between those who have essentially unlimited time and those who always race against the clock to make ends meet and stay alive, living literally day by day.

Seeing currency and, therefore, wealth as a direct measure of someone's life expectancy is a clear thought experiment. Much like the inflation in so many works of science fiction, prices for consumers in *In Time* rise faster than wages, keeping the working class in poverty.

Inflation comes up repeatedly in science fiction. It is no coincidence: many authors have witnessed firsthand how inflation decreases purchasing power and harms quality of life. Yet governments welcome at least some inflation, printing money to appease various constituents and special-interest groups. Inflation is effectively a hidden tax, as it reduces the relative wealth of consumers through price increases while devaluing the government's debt. It is also a regressive tax, as the wealthy invest in assets that outperform inflation while the working class spends more of its income on now higher-priced goods.[10]

24.4.2 Cash and Debt

As technology has become more pervasive, more transactions have become cashless. Dan Simmons creates an interstellar civilization in *Hyperion*, which necessitates more electronic transacting: "If our society ever opted for Orwell's Big Brother approach, the instrument of choice for oppression would have to be the credit wake. In a totally noncash economy with only a vestigial barter black market, a person's activities could be tracked in real time by monitoring the credit wake of his or her universal card. There were strict laws protecting card privacy but laws had a bad habit of being ignored or abrogated when societal push came to totalitarian shove" (Simmons 1991, p. 360). As Simmons warns, control over spending can lead to surveillance and invasion of privacy. While physical money carries its own risks such as theft, digital money depends on a robust, secure infrastructure.

Cryogenic freezing and near-lightspeed travel both create scenarios in which people can effectively travel faster into the future without aging. Because some investments tend to generate returns that exceed inflation over time, jumping into the future could potentially allow an investor to benefit sooner in his own life from the returns. Unfortunately, unmanaged investments can go awry: "Mother sent me to this back end of the outback on a Phase Three ramship, *slower than light*, frozen with the cattle embryos and orange juice concentrate and feeder viruses, on a trip that took *one hundred and twenty-nine* shipboard years, with an objective time-debt of *one hundred and sixty-seven standard years*! Mother figured the accrued interest on the long-term accounts would be enough to pay off our family debt and perhaps allow me to survive comfortably for a while. For the first and last time in her life, Mother figured wrong" (Simmons 1991, p. 196). Consequently, the narrator arrives on a new planet completely destitute.

Governments, companies, and individuals often borrow capital to invest, hoping that the real return will exceed the interest of the loan. While this can serve an actively managed company (where the borrower controls how the capital is deployed), that is not the case for passive investments, especially when the investor is literally incapacitated. There are always risks, including "black swan" events that no model can predict.

24.4.3 Value of Property

One narrator in *Hyperion* ultimately becomes very wealthy: "My home has thirty-eight rooms on thirty-six worlds [thanks to portals between worlds]. … the guest bathroom … consists of toilet, bidet, sink and shower stall on an open, wall-less raft afloat on the violet seaworld of Mare Infinitus" (Simmons 1991, pp. 205–206). Why would someone want his house to span across dozens of worlds? Technology is often first sold to early adopters with high disposable income, whose motivation might be of pure vanity. But without those customers, innovation could not generate scale economies.

In Philip José Farmer's *To Your Scattered Bodies Go*, all humans throughout history are resurrected along a river. Necessities and comforts are provided, resulting in immediate trade among people who value their provided goods differently. A reincarnated Sir Richard Burton seeks to travel up the river. Of course, he must build a boat himself: "Burton could build a craft. However, the people hereabouts were conservationists. They did not believe in despoiling the land of its trees. Oak and pine were to be left untouched, but bamboo was available. Even this material would have to be purchased with cigarettes and liquor, which would take some time to accumulate from his grail" (p. 212). Burton can choose to trade away his goods over time to acquire the bamboo, which he considers more valuable. The locals have set their own values, unwilling to accept any price for oak and pine trees.

In Roger Zelazny's novel *This Immortal*, aliens attempt to appropriate the Earth's wonders and turn them into tourist attractions. The protagonist develops a strategy to dissuade the aliens: "'Conrad [the alien asks him], why are you tearing down the pyramid?' … 'To let you know if you want this place and you do manage to take it away from us, you'll get it in worse shape.'" (Zelazny 1966, p. 161). Conrad, then, has set up a kind of "poison pill" similar to that used to discourage corporate raiders. In the modern world, such shareholder-rights plans help protect companies from hostile takeovers. Of course, Conrad is not just devaluing shares in an asset by creating new ones at a discount; instead, he destroys that asset, in this case an ancient and irreplaceable pyramid.

In 2021, the Chinese government threatened and placed restrictions on technology companies whose publicly traded shares were listed in the United States. Consequently, the value of those companies fell. These actions harmed Chinese citizens as well as foreign investors, but they may have the ultimate effect (intended or not) of allowing more Chinese investors to benefit from the appreciation of Chinese companies.

In Anne McCaffrey's *Dragonflight*, the protagonist intentionally sabotages her family's fortress after it is stolen in order to force the usurper to give it up. As the usurper says, "The day one of my Holds cannot support itself *or* the visit of its rightful overlord, I shall renounce it" (McCaffrey 1996, p. 37). Property is not valuable in and of itself. It often comes with maintenance costs, which can lead to the abandonment of the asset over time when it becomes uneconomical to manage.

24.4.4 Resources

In the film *Interstellar*, directed by Christopher Nolan, climate change has drastically reduced the ability to grow crops, thereby risking human extinction. Humanity must find a way to travel to a wormhole that has inexplicably opened near Saturn and then travel to a distant galaxy with habitable worlds. Such a solution to a climate crisis is extremely unlikely; the movie warns viewers to protect Earth's long-term health.

It is difficult to plan for resource use in the long term. Given technology's rapid development, is it worth harming short-term economic growth to avert a potential future catastrophe that might be easier to prevent later? With multiple countries involved, a short-term sacrifice could result in a disadvantage absent international cooperation.

Human capital is also a resource, particularly when individuals have irreplaceable expertise and proficiencies; skills that require immense training make especially valuable those who have the skills. Orson Scott Card's military science-fiction novel *Ender's Game* explores this concept: "Human beings are free except when humanity needs them. Maybe humanity needs you. To do something. Maybe humanity needs me—to find out what you're good for. We might both do despicable things, Ender, but if humankind survives, then we were good tools" (Card 2002 p. 35). Ender himself is the resource, as much as the materials used in the ships and the weapons deployed in the conflict.

In the dystopian film *Mad Max: Fury Road*, directed by George Miller, warfare ensues over the limited remaining manufactured resources: gasoline and ammunition. While this scenario is extreme—with no new supply—it shows how dependent civilization can become on certain manufactured and processed goods. The conflict over gasoline mirrors the oil crises of the 1970s, in which the Organization of the Petroleum

Exporting Countries (OPEC) declared an embargo in 1973 and then the oil production of Iran and Iraq declined drastically in 1979–1980. Even more-recent shocks in energy prices have had profound effects. Although, thanks to fracking and other innovations, the United States can produce enough oil and natural gas to sustain itself, there is no instantaneous on–off switch for adjusting that production. As global prices fall, for example, it becomes uneconomical for domestic oil to compete with cheap imports. Also, government regulation—or the specter of it—can discourage production.

Modern electronic goods are dependent on semiconductor chips. The COVID-19 pandemic halted production at many factories, with widespread expectations of a decrease in demand. Instead, an economy adapting to remote work *increased* demand for various commodities, including canned goods and consumer-sized paper goods. Shortages affected a wide range of goods, including less obvious ones such as automobiles (for which electronics are a small but crucial component).

24.4.5 Productivity

While *Mad Max* occurs in a world with more destruction than productivity, Ayn Rand's novel *Atlas Shrugged* explores a productive society in which private enterprises suffer under increased governmental regulation.

Unlike many other authors' works, which depict workers exploited by capitalism, Rand views capitalists as exploited by government. Rather than the government spending money, she believes those who have generated capital are best equipped to deploy it: "Let me give you a tip on a clue to men's characters: the man who damns money has obtained it dishonorably; the man who respects it has earned it" (Rand 1957, p. 383). Moreover, she does not see capital as an end, but rather a means: "But money is only a tool. It will take you wherever you wish, but it will not replace you as the driver" (Rand 1957, p. 382). Governments often throw money at problems, as if money were the direct solution.

Rand sees productivity arising from individuals, not from central planning: "Productiveness is your acceptance of morality, your recognition of the fact that you choose to live—that productive work is the process by which man's consciousness controls his existence, a constant process of acquiring knowledge and shaping matter to fit one's purpose, of translating an idea into physical form, of remaking the earth in the image of one's values—that all work is creative work if done by a thinking mind … that your work is yours to choose, and the choice is as wide as your mind … that your work is the process of achieving your values" (Rand 1957, p. 933). Rand focuses on the individual over the state. As both history and science fiction have demonstrated, governments do not necessarily survive forever.

While Rand's philosophy is clearly conservative, more-liberal authors such as Jemisin appear to agree that pursuing one's own passions is preferable to following societal norms.

24.5 Trade

Rand's economic vision of an individual-focused economy can work because individuals find niches in society and trade with each other to provide for their needs and desires. Individuals establish value for their work through transactions, allowing them to make decisions that enhance their productivity and income. Currencies provide an intermediary store of value so that bartering does not have to be done directly between what individuals produce for society and what they consume.

24.5.1 Monetary System

Without a monetary system, in which government standardizes a currency for commerce, people must barter for goods. This old practice occurs in Brian Aldiss's novel *Starship*, where civilization on a generation ship—in which the journey will span multiple generations—has degenerated into a primitive society:

> This deteriorating state of affairs Complain simply attributed to a grudge Roffery the Valuer held against the hunter clan, being unable to integrate the lower prices Roffery allowed for wild meat with the abundance of domestic fare. Consequently, he pushed through the market crowd and greeted the valuer in surly fashion.
>
> "'spansion to your ego," he said grudgingly.
>
> "Your expense," the Valuer replied genially, looking up from an immense list he was painfully compiling. "Running meat's down today, hunter. It'll take a good sized carcass to earn six loaves."
>
> "Hem's guts! And you told me wheat was down the last time I saw you, you twisting rogue."

> "Keep a civil turn of phrase, Complain: your own carcass isn't worth a crust to me. So I did tell you wheat was down. It is down—but running meat's down more". (Aldiss 1969, pp. 19–20)

Aldiss portrays a society in which there is no currency at all; thus, a "Valuer" must compile a tremendously extensive list stating, for a given commodity, how much of each other commodity it is worth. That is, every commodity has a price in terms of every other commodity—and its price is in physical units, as money does not exist.

Despite the absence of money, the characters still have a concept of price levels—a kind of "ghost money." From a historical perspective, there was a period before money existed when market transactions still occurred. Aldiss imagines a scenario in which a future human society regresses back to these barter transactions!

Even when currency does exist, people can nevertheless barter and transact—using other intermediary goods. In the film *District 9*, directed by Neil Blomkamp, alien weapons—as well as terrestrial ones—are used as currency. Any good can act as a store of value. Guns, in fact, are more immediately practical than metals.

Historically, both gold and silver were used as stores of values prior to fiat currency, and currencies were backed by precious-metal reserves. In the film *Looper*, directed by Rian Johnson, assassins kill victims from the future sent to the past by an organized crime syndicate. They are paid with silver bars, a standardized store of value across both time and space.

What happens in a world where metals are so rare that technology is adversely affected? Robert Silverberg's novel *Lord Valentine's Castle* takes place on a massive planet with minimal natural metals. Consequently, the world and its economy feel mostly ancient, and there is no means of electronic payment. As the currency is explained, "These sausages cost ten weights. A hundred weights make a crown, ten crowns make a royal, and this [coin] is fifty of those" (Silverberg 1981, p. 10). The flexibility of coins of various values allows commerce to occur with a high level of trust.

Metals in *Lord Valentine's Castle* are hoarded for pretentious purposes: "It [the performers' stage] all floats on a pool of quicksilver … You could buy three provinces with the value of the metal" (Silverberg 1981, p. 112). Of course, if that amount of silver entered the marketplace, it could very well collapse the price. Who would be willing to buy that much of it in a world with no capacity to manufacture it into something useful?

Yet even today people do store some of their wealth in precious metals. Gold carries tremendous cultural value in countries such as India, and gold is a vehicle of speculation for investors worldwide.

24.5.2 Commerce

Walter M. Miller's *A Canticle for Leibowitz* follows events at a Catholic monastery protecting scientific knowledge after a devastating nuclear war. Because the monks cannot judge which documents are valuable, much of the so-called Memorabilia seems frivolous to a contemporary reader. The value of the Memorabilia changes over time. At first, it is completely useless because society lacks the capacity to utilize it. But as technology is rebirthed, some of it proves extremely valuable for a time: "'Since the death of the last civilization, the Memorabilia has been our special province, Benjamin. And we've kept it. But now? I sense the predicament of the shoemaker who tries to sell shoes in a village of shoemakers.' … 'It could be done, if he manufactures a special and superior type of shoe'" (Miller 1961, p. 143).

Rare goods command high value due to lack of supply. But as time passes and more competition occurs—in this case, from other inventors gaining the resources to rebuild—these same assets become less valuable. As certain goods become commoditized, offering a premium product becomes one way to justify a premium price.

Even a standardized commodity can become valuable given the right supply and demand circumstances. In *Dune*, Frank Herbert introduces a rare material found only on one planet, where it is guarded by giant sandworms: "By giving me Arrakis, His Majesty is forced to give us a CHOAM [company name] directorship … [The spice] cannot be manufactured, it must be mined on Arrakis…Imagine what would happen if something should reduce spice production" (Herbert 1990, pp. 42–43). Civilization is dependent on a drug ("the spice") to enhance awareness, allowing navigators to chart safe routes through space–time and enable practical interstellar travel.

Certain goods are essential to an economy during given periods of time. Thus, monopolies of supply can lead to skyrocketing prices, as again evidenced by the 1970s oil crises with OPEC controlling supply.

Other works of science fiction paint an interstellar civilization without the ability to break the speed of light in travel or communication. Thus, governments cannot practically govern across light-years, as their agents and the general population are unlikely to act in the interests of the home world when the colony is so remote, far more distant than America from Europe during the colonial period.

As Earth's civilizations developed, maintaining sprawling empires was impossible. Many European colonies gained independence. Even Europeans inserted in colonial governance roles have acted outside the interests of their rulers. Christopher Columbus, for instance, was removed as governor of the West Indies.

Taking trade into the interstellar context addresses the costs and risks in moving goods across distances. Although transportation costs around Earth are more reasonable, there are risks to relying on international supply chains—as the United States and China have recently discovered.

In Vernor Vinge's *A Deepness in the Sky*, a group of traders called the Qeng Ho arrives at a world hoping to make a deal for ancient alien technology: "[T]here was the possibility of trade. Here, well, there was treasure but it did not belong to either side. It lay frozen, waiting to be looted or exploited or developed, depending on one's nature" (Vinge 2000, p. 19). Trade, not government treaties, tends to unite people across distinct cultures. As Vinge notes, "No government can maintain itself across light-years. Hell, most governments don't last a few centuries. Politics may come and go, but trade goes on forever" (Vinge 2000, p. 222). In a futuristic context, government could not practically rule across light-years of distance. Both communication and travel would take too much time. Only economic incentives could unite humanity across the galaxy.

While the original explorers of the Americas were funded by European states, they were motivated more by trade or plunder than politics or ideology. Even amidst revolutions and coups, companies have survived. Ultimately, people do not work only for government; they prioritize their own interests. A nod to Ayn Rand!

24.5.3 Principal-Agent Problem

In recent years, the sharing economy has grown, adding new value to existing capital such as homes and vehicles by increasing their use. One complication is that lessees or renters do not fully own that capital, leading to potential damage emanating from the inherent principal-agent problem. Simple solutions like insurance and background checks are used, but they are not foolproof.

In *Mindswap*, Robert Sheckley examines the principal-agent problem of swapping bodies, a different kind of vacation rental:

> "Next, you and the Martian Gentleman will both sign a Reciprocal Damage Clause. This states that any damage to your host body, whether by omission or commission, and including Acts of God, will, one, be recompensed at the rate established by interstellar convention, and, two, that such damage will be visited reciprocally upon your own body in accordance with the *lex talionis*."
>
> "Huh?" Marvin said.
>
> "Eye for eye, tooth for tooth," Mr. Blanders explained. "It's really quite simple enough. Suppose you, in the Martian corpus, break a leg on the last day of Occupancy. You suffer the pain, to be sure, but not the subsequent inconvenience, which you avoid by returning to your own undamaged body. But this is not equitable. Why should you escape the consequences of your own accident? Why should someone else suffer those consequences for you? So, in the interests of justice, interstellar law requires that, upon reoccupying your own body, your own leg be broken in as scientific and painless a manner as possible."
>
> "Even if the first broken leg was an accident?"
>
> "*Especially* if it were an accident. We have found that the Reciprocal Damage Clause has cut down the number of such accidents quite considerably". (Sheckley 1966, p. 17)

Being accountable will change individuals' actions. Accidents are often avoidable through additional defensive action. While lessees and renters may not have legal liability for damage caused by others, they do have some ability to deter such damage.

When it comes to governance, both political and corporate, managers' own interests do not always align with those of their constituents, shareholders, or employees. Leaders such as Egypt's Hosni Mubarak, Iraq's Saddam Hussein, Libya's Muammar Gaddafi, and Russia's Vladimir Putin have been accused of such corruption, amassing vast personal fortunes at the expense of their constituents.

24.5.4 Intellectual Property

Intellectual property is increasingly a source of capital in the modern world. Corporate lawsuits are often about this kind of theft. Not only is intellectual property easier to steal, but also it is simpler to use than physical materials, making it extremely valuable.

Neal Stephenson forecasted the exploding value of intellectual property in the computer age in his 1992 novel *Snow Crash*: “[W]hen I have a programmer working under me who is working with that information, he is wielding enormous power. Information is going into his brain. And it’s staying there. … he doesn’t have any right to that information. If I was running a car factory, I wouldn’t let workers drive the cars home or borrow tools. But that’s what I do at five o’clock each day, all over the world, when my hackers go home from work” (Stephenson 2003, p. 116). Companies have increasingly become valued not on current revenue but on a projection of future growth. And future growth comes in large part from intangibles such as intellectual property and customer loyalty.

For intellectual property to be recognized, especially due to its intangible nature, it must be recognized by a government through polices such as the issuance of patents. Because the government grants property rights, it is possible for the government to claim ownership of all ideas. In Ursula K. Le Guin’s *The Dispossessed*, the protagonist says, “I didn’t understand that here an idea is a property of the State” (Le Guin 2006, p. 293). The researcher ultimately refuses to share his work with the other theoreticians in that society, for he fears it would be squandered. Thus, government control can suppress innovation.

24.6 Government

24.6.1 Central Planning

While Ayn Rand viewed government as inherently burdensome, other authors have exposed flaws in specific systems of governing. In *Nineteen Eighty-Four*, George Orwell warns of “Big Brother” totalitarian governance, as well as dividing the world into a few superstates—which he feared was occurring after World War II. Indeed, the Cold War resulted from the geopolitical influence of Soviet communism and American democracy after the war. Orwell’s fictional superstates end up in a perpetual war: "The economy of many countries was allowed to stagnate, land went out of cultivation, capital equipment was not added to, great blocks of the population were prevented from working and kept half alive by State charity. But this, too, entailed military weakness, and since the privations it inflicted were obviously unnecessary, it made opposition inevitable. The problem was how to keep the wheels of industry turning without increasing the real wealth of the world. Goods must be produced, but they must not be distributed. And in practice the only way of achieving this was by continuous warfare" (Orwell 1981, p. 155). The fictional nations use war to keep their economies in a constant state of absorbing goods before they can benefit the lower classes, while providing some basic support to them. Such a system maintains the status quo and prevents any social mobility by forcing dependence on the state.

The government effectively freezes history: "Every record has been destroyed or falsified, every book rewritten, every picture has been repainted, every statue and street building has been renamed, every date has been altered. And the process is continuing day by day and minute by minute. History has stopped. Nothing exists except an endless present in which the Party is always right" (Orwell 1981, p. 127). Put more succinctly, "'Who controls the past,' ran the Party slogan, 'controls the future: who controls the present controls the past'" (Orwell 1981, p. 31). Through actions such as war and propaganda, a government can unite its citizens, even against their own interests.

In *Brave New World*, Aldous Huxley imagines a command economy focused on mass production and consumerism, where even the people are themselves manufactured. Instead of ancestral castes, citizens are engineered: "But though the Epsilon mind was mature at ten, the Epsilon body was not fit to work till eighteen. Long years of superfluous and wasted immaturity. If the physical development could be speeded up till it was as quick, say, as a cow's, what an enormous saving to the Community!" (Huxley 1998, p. 15).

For consumption to increase, consumers cannot spend their time on pursuits that do not require the manufacture of goods: "A love of nature keeps no factories busy" (Huxley 1998, p. 23). The economy of Huxley's world, inspired in part by Henry Ford's assembly line, bases its strength on consumption, as "that's the price we have to pay for stability. You've got to choose between happiness and what people used to call high art. We've sacrificed the high art" (Huxley 1998, p. 220). Society, then, abandons individual interests and pursuits. Instead, everyone's time is spent on production and consumption of the goods and services approved by the state. But it is not only art that suffers: “Science is dangerous; we have to keep it most carefully chained and muzzled” (Huxley 1998, p. 225). Changes in technology would inevitably disrupt the status quo.

In both dystopian works, production is organized by the state rather than by individuals. There are certain injustices and inefficiencies which private industry will not necessarily address; absent regulation, it is not in a company’s best interest to manage pollution. That is where government comes in. But as Orwell, Huxley, and especially Rand have warned: when the government overmanages, citizens lose their freedom, harming both culture and the economy.

24.6.2 Property Law

For property rights to exist, they must be recognized by government. Without some system to verify and recognize property, anyone could claim ownership of anything, leading to endless disputes. Indeed, property owners historically *were* the government (feudalism) or directly elected the government (early America).

In *Stranger in a Strange Land*, Robert A. Heinlein depicts a human raised by Martians as he visits Earth for the first time. Heinlein imagines a drought-stricken planet in which the Martians must “share water and grow closer” (Heinlein 1987, p. 289). As a result of the planet’s environment, the “Martians seemed to have defeated death, and they seemed not to have money, property, nor government in any human sense” (Heinlein 1987, p. 142). The protagonist, then, views property far differently from Earthlings: “Nor do I regard that wealth as ‘his’; he didn’t produce it. Even if he had earned it, ‘property’ is not the natural and obvious concept that most people think it is” (Heinlein 1987, p. 185). Although Heinlein demonstrated clear conservative views in his fiction, he did value aspects of fictional societies that did not match his own politics.

As the protagonist impacts Earth, he has no intention of eliminating property rights: “No, money and property will not disappear—Michael says that both concepts are useful—but they’re going to be turned upside down and people will have to learn new rules (the hard way, just as we have) or be hopelessly outclassed. What happens to Lunar Enterprises when the common carrier between here and Luna City is teleportation?” (Heinlein 1987, p. 400). New technologies and other changes can significantly alter the value of certain property.

In *The Dispossesed*, the protagonist travels from a communist world to a capitalist society on a twin planet. There, he sees that capitalism "made the superb cars and comfortable trains. The lure and compulsion of profit was evidently a much more effective replacement of the natural initiative than he had been led to believe. … [He] could see how efficiently a propertarian economy ran its manufacturing and power supply" (Le Guin 2006, pp. 82–83). Capitalism clearly promotes innovation and production, as it rewards individuals for their contributions, allowing the best ideas—rather than those the government or culture chooses—to succeed.

With a completely "flat" society, "[w]e have no states, no nations, no presidents, no premiers, no chiefs, no generals, no bosses, no bankers, no landlords, no wages, no charity, no police, no soldiers, no wars. Nor do we have much else. We are sharers, not owners. We are not prosperous. None of us is rich. None of us is powerful" (Le Guin 2006, p. 300).

However, promoting equity, in which outcomes are the same, often means harming those who might succeed. Progress is stifled. In *Children of Men*, James notes that "equality is a political theory not a practical policy" (James 2006, p. 7). Allowing individuals to succeed beyond their peers can lift all of society, even when it allows inequality. This is the classic trade-off between equity and efficiency.

Back on the communist world of *The Dispossessed*, "[m]ost refectories served dessert once or twice a decad [ten-day week]. Here it was served nightly. Why? Were the members of the Central Institute of the Sciences better than other people? … He weighed the moral discomfort against the practical advantage, and found the latter heavier. … The responsibility justified the privilege" (Le Guin 2006, pp. 111–112). Ideas of communism quickly break down as a certain privileged class justifies greater and greater rewards. Do those who work more deserve more food? This concept evolves for the protagonist, as he sees the value in an "economy based on the principle that each worker is paid as he deserves, for the value of his labor—not by capitalists whom he's forced to serve, but by the state of which he's a member!" (Le Guin 2006, p. 135). Here, he seems to shift to believing not that a pure free market can provide just compensation—he sees that often it does not—but that a market with some government intervention can do so.

On the capitalist world, he observes, "The rich are very rich indeed, but the poor are not so very poor. They are neither enslaved nor starving" (Le Guin 2006, p. 341). Modern China did not achieve explosive economic growth until the communist government allowed individuals to prosper, including the development of a middle class. While capitalism has flaws, its incentives match human nature.

Capitalism allows the ambitious to prosper, but it will not provide a safety net in and of itself. The protagonist observes this dichotomy: “Because there is nothing here but States and their weapons, the rich and their lies, and the poor and their misery. There is no way to act rightly, with a clear heart, on Urras. There is nothing you can do that profit does not enter into, and fear of loss, and the wish for power. You cannot say good morning without knowing which of you is ‘superior’ to the other, or trying to prove it. You cannot act like a brother to other people, you must manipulate them, or command them, or obey them, or trick them. … I know it’s full of evils, full of human injustice, greed, folly, waste. But it is also full of good, of beauty, vitality, achievement. It is what a world should be! It is alive, tremendously alive—alive, despite all its evils, with hope” (Le Guin 2006, pp. 345–346).

Le Guin does not take a side on the political debate. Instead, she highlights the flaws of the various systems. But she shows that deviating slightly from the rigid theories—allowing some government intervention into the otherwise free market—can work under certain circumstances.

24.6.3 Capitalism, Socialism, and Communism

Le Guin surmises that a harsh environment necessitates a communal society to survive: “This planet wasn’t meant to support civilization. If we let one another down, if we don’t give up our personal desires to the common good, nothing, nothing on this barren world can save us. Human solidarity is our only resource” (Le Guin 2006, p. 167). This conclusion is reasonable; given such a challenge, the circumstance could be overcome only by cooperation.

Robert A. Heinlein, on the other hand, takes the opposite approach. In *The Moon is a Harsh Mistress*, a more-anarchistic society emerges on a lunar penal colony. Under such harsh conditions, Heinlein supposes that only a society that rewards individual behavior could thrive. Thus, he uses this setting to promote libertarian ideals: “There is no worse tyranny than to force a man to pay for what he does not want merely because you think it would be good for him” (Heinlein 2018, p. 308). When resources are limited and survival is difficult, taxes deprive individuals from making their own decisions. Instead, the government spends in ways that might benefit some citizens over others.

Of course, what if the economy is poorly managed? “But the solution is so simple that you all know it. Here in Luna we’re rich. Three million hardworking, smart, skilled people, enough water, plenty of everything, endless power, endless cubic. But... what we don’t have is a free market. … Authority charges too much for water, don’t buy. It pays too little for ice, don’t sell. It holds monopoly on export, don’t export. Down in [Mumbai] they want wheat. If it doesn’t arrive, the day will come when brokers come here to bid for it—at triple or more the present prices!” (Heinlein 2018, p. 22). Leadership does not always act in the best interest of its stakeholders. If leaders act for their own personal gain, they may make deals that harm those whom they represent, for example, taking kickbacks in exchange for an uneconomical deal.

Even in a democracy, certain classes form: “A managed democracy is a wonderful thing … for the managers … and its greatest strength is a ‘free press’ when ‘free’ is defined as ‘responsible’ and the managers define what is ‘irresponsible’” (Heinlein 2018, p. 259). The political class is perceived to be hypocritical and entitled. In our world, elected and appointed politicians often engage in business with lobbyists after stepping down from government.

Even though governments uphold laws, individuals do not always follow them: “I will accept any rules that *you* feel necessary to *your* freedom. *I* am free, no matter what rules surround me. If I find them tolerable, I tolerate them; if I find them too obnoxious, I break them. I am free because I know that I *alone* am morally responsible for everything I do” (Heinlein 2018, p. 78). Some of the modern economy occurs outside of government oversight. People transact in cash or cryptocurrency, avoiding sales and income taxes and engaging in illegal transactions.

Philip José Farmer’s novella *Riders of the Purple Wage* creates a setting closer to the modern world, in which all citizens receive a universal basic income—a traditional political goal of progressives. As technology develops the means to automate the production of necessities, a human labor force may not be needed to provide basic goods: “There is no more starvation or want anywhere, except among the self-exiles wandering in the woods. And the food and goods are shipped to the pandoras and dispensed to the receivers of the purple wage. The purple wage. A Madison-Avenue euphemism with connotations of royalty and divine right. Earned by just being born” (Farmer 1972, p. 89).

By providing these goods to consumers, the government can insert itself into the most basic transactions. Money would not be needed: "Actually, the money has no value now except as collector's items. Shortly after the theft, the government called in all currency and then issued new bills that could not be mistaken for the old" (Farmer 1972, p. 128). By controlling currency, governments can monitor the activities of its citizens. Farmer wrote *Riders of the Purple Wage* before credit cards were commonplace (allowing easier surveillance); but, with digital currencies, electronic transfer, and other financial innovations, the possibility of physical currency becoming obsolete is real.

On regulation, Farmer describes a gray area thanks to corruption: "The [government] has no overt objection to privately owned taverns, run by citizens who have paid all license fees, passed all examinations, posted all bonds, and bribed the local politicians and police chief. Since there is no provision made for them, no large buildings available for rent, the taverns are in the homes of the owners themselves" (Farmer 1972, p. 95). People act in their own self-interest, not at the whims of government. Central planning goes against human nature. A system that embraces human nature might do better than one that tries to suppress it. Again, a nod to Ayn Rand!

By crushing innovation, the government dooms the general population to survive on welfare, while the political class can remain comfortable and in power: "The officials and workers were getting relatively high wages, but many citizens had to be contented with their guaranteed income" (Farmer 1972, p. 128). Equality (among the masses), then, is achieved not by raising people up, but by keeping them down.

In some localities (Chicago, New York, etc.), public employees earn more than the average constituent, who pays the taxes that support them. Public unions support political candidates who in turn agree to employment terms that benefit the union members. Without judging whether this is right, there is certainly the risk of a two-tiered economy in which public employees enjoy greater prosperity at the expense of working-class taxpayers.

Note that consumption-based economies are not tied to capitalism. Even a socialist welfare state like Farmer's is based on consuming goods to appease the masses.

In an early work of science fiction from 1895, H. G. Wells warns of the dangers of *both* capitalism and communism. The Eloi society the time traveler first experiences seems like a communist utopia, but he eventually learns the truth: "The too-perfect security of the Upper-worlders had led them to a slow movement of degeneration, to a general dwindling in size, strength, and intelligence" (Wells 1996, p. 49). The time traveler learns that the Eloi people have a symbiotic relationship with another people, the Morlocks. These workers manufacture goods for the Eloi, who are then sacrificed as food for the Morlocks. The allegory epitomizes the dangers of capitalism—a working class doomed to serve machines metaphorically and literally underground—while also showing how a communist society can stifle innovation and lead to societal regression.

Wells does not seek to convince the reader how to judge societies, or of what system works best. Instead, he highlights the dangers and pitfalls. Communist ideals often oversimplify human nature, just as *laissez-faire* capitalism leaves individuals and communities vulnerable to economic shifts and negative externalities. Fortunately, economic systems can be tweaked; fiction, in this case, provides some warnings to identify these nuances.

24.7 Alternate History

History is not an experimental science. Historical data cannot as easily expose flaws as can fiction. That is because history is limited to what has already happened.

Science fiction often modifies one technology or event outside the world's *current* trajectory and then extrapolates from that. Alternate history does something similar, but from the distant or not-so-distant *past* into the future.

Some alternate history might result from changes in technology, altering timelines or the people involved. Other counterfactual history can result from reasonable changes based on probability. Outcomes are not deterministic, and small random factors can impact history.

24.7.1 Economic Cost (and Benefit) of War

Outcomes of war represent major historical events; thus, swapping the winners and losers makes for interesting fiction. Technology has played a pivotal role in multiple wars, with the historical ultimate being the atomic bomb in World War II. This technology extended past the war and into the development of nuclear power.

In *The Guns of the South*, Harry Turtledove gives the South a technological advantage in the Civil War. Consequently, slavery was not abolished as early: "The Confederate constitution enshrined the right to own slaves and trade them within the nation's borders. The Southern economy rested on the backs of its black labor force. But a lot of people who could never have stomached the butcher's trade ate meat" (Turtledove 1993, p. 199). Historically, while slavery was prevalent in America, it also benefited European colonial powers directly. For example, Belgium exploited enslaved persons in the Congo, out of sight of its domestic citizens.

Ward Moore's short story *Bring the Jubilee* occurs in an alternate timeline in which the South similarly won the Civil War. In Moore's story, the North and South become two disparate nations, with the North struggling to recover: "But the Peace of Richmond had also laid the cost of the war on the beaten North … The postwar inflation entered the galloping stage … and precipitated the food riots of 1873 and '74" (Moore 2001, p. 153). With the antebellum North's economy centered on industry, while the South's on agriculture, the North had to rely on imports for basic survival, reducing the value of its own goods and services. Once again, an author presents a cautionary tale of inflation.

Although the North specializes in manufacturing, "the great issue in [the North's] Congress was the never-completed Pacific transcontinental line, though Canada had one and the Confederate States seven" (Moore 2001, p. 155). An economy cannot function well under uncertainty. The Civil War's actual outcome was necessary to prevent the "pervasive fear of imminent war" (Moore 2001, p. 172) that would have stopped progress. The alternative would have been a two-state solution, with the threat of war always imminent. A unified America helped usher in prosperity.

Indeed, as the protagonist ultimately travels back in time, he observes that "[a]ny one of the inventions of my own time would make me a rich man if I could reproduce them" (Moore 2001, p. 247). A lot of value was created through continued innovation after the Civil War. Would it have still occurred as quickly if the threat of war had dragged on?

Moore also presents the oft-forgotten issue of indentured servitude: "Indenting's pretty strictly regulated. That's the idea, anyway. You can't be made to work over 60 hours a week—ten hours a day. With $1,000 or $1,200 you could get all the education you want in your spare time and then turn your learning to account by making enough money to buy yourself free" (Moore 2001, p. 165).

Although slavery and indentured servitude might appear to provide cheap labor that would increase overall production, they force people into roles not based on their talents or motivation, which cannot help productivity.[11] Slavery limits collaboration and becomes little different from a centrally planned economy: the slaveowners rather than government make the decisions regarding work.

Other wars have had similarly large impacts on the United States and its economy. Philip K. Dick paints an alternate history in *The Man in the High Castle*, in which the Axis powers won World War II, dividing the United States between Japan and Germany. In this scenario, much of the United States was destroyed, as Europe had been in reality: "This is just the sticks to you, the Rockies. Nothing has happened here since before the war. Retired old people, farmers, the stupid, slow, poor … and all the smart boys have flocked east to New York, crossed the border legally or illegally. Because, she thought, that's where the money is, the big industrial money. The expansion. German investment has done a lot … it didn't take long for them to build the U.S. back up" (Dick 1988, p. 30). With America divided, and the Rockies a neutral zone, opportunity exists only on the coasts, thanks to trade and capital flows.

In reality, the continental United States was untouched in the war. The United States then benefited from rebuilding Europe, just as fictional German investment did in Dick's work. And instead of the United States being divided, Germany was divided. Dick also references how individuals sought opportunity across borders. In the same way, many people in East Germany illegally crossed (or attempted to cross) the Berlin wall.

24.7.2 Historical Impact of Government

As governments have changed, they have often contributed to instability and unpredictability. A change in rule, as often occurred during the Middle Ages, impacted everyone from vassals to lords.

In *The Alteration*, Kingsley Amis imagines history without the Reformation. With governments continuing to answer to the Catholic Church, innovation became stifled: "[Diesel] ignition was achieved merely by compressing petroleum vapour to a certain density, without the introduction of a spark. … electricity was appallingly dangerous, both as it existed and as it might be developed. No wonder its exploration had never received official encouragement, nor that persistent rumours told of such exploration by inventors in New England" (Amis 1976, p. 8).

When governments discourage certain technologies and enforce arbitrary laws, the economy undoubtedly suffers.

In the case of the Church, there is plenty of historical evidence to suggest Amis's fiction could have been a reality under different circumstances. Galileo Galilei spent the end of his life under house arrest after being tried by the Roman Inquisition, which also banned books by scientists such as Nicolaus Copernicus. Had the Catholic Church grown in its power, rather than losing some to Protestantism, such arrests and bans would have curbed innovation. Because technology drives economic growth, Europe would have fallen economically behind the rest of the world.

Kim Stanley Robinson's *The Years of Rice and Salt* creates an extreme alternate history, in which the Black Death wiped out nearly all of Europe's population rather than only the historical one-third, enabling other peoples of the world to spread into Europe and the Americas. Islam would have become a larger force globally, as more nations would have been majority Muslim. Of course, Islam is analogous to the broad religion of Christianity: just as Catholicism does not represent all of Christianity (unlike in *The Alteration*, of course), Islam is itself divided into factions such as Sunni and Shia. For the purposes of diplomacy, however, other nations might consider all those nations a singular bloc:

> The whole of Islam was accused of breaking the commitments forced on them at the Shanghai Conference after the war, as if Islam were a monolithic block, a laughable concept even in the depths of the war itself. Sanctions and even embargoes were being called for in China and India and Yingzhou. The effect of the threat alone was felt immediately in Firanja: the price of rice shot up, then the price of potatoes and maple syrup, and coffee beans. Hoarding quickly followed, old wartime habits kicking in, and even as prices rose staples were cleared off the shelves of the groceries the moment they appeared. This affected everything else as well, both food and other matters. Hoarding was a very contagious phenomenon, a bad mentality, a loss of faith in the system's ability to keep everything running; and as the system had indeed broken down so disastrously at the end of the war, a lot of people were prone to hoard at the first hint of a scare" (Robinson 2002, p. 556).

In this case, the actions of foreign governments have immense impact on the Muslim world. Shortages do lead to further shortages due to hoarding (example: some commodities during the COVID-19 pandemic), as individuals make rational decisions based on past experience. Threats of international action, such as the proposed embargoes in Robinson's novel, are enough to move markets.

Comparable actions have affected actual history. OPEC has historically represented most international oil exports. Consequently, OPEC can control prices by adjusting output. But not all its members follow agreements to restrict production. The threat is often enough to have the desired effect on prices.

The television show *For All Mankind*, produced by Ronald D. Moore, imagines the past and future of space exploration if the Soviet Union had landed on the Moon prior to the United States. This outcome would have prolonged the space race and had a variety of social and economic consequences. In the show, these changes accelerate the civil-rights movement while advancing the technical capabilities of both the Americans and Soviets. From an economic perspective, what would have happened if the United States government had continued spending significant resources on space exploration? The benefits of such spending are difficult to predict; scientific breakthroughs can benefit the entire terrestrial economy. The Apollo missions, for instance, led to advances in computing technology used in virtually every industry today.

24.7.3 Parallel Worlds

Larry Niven's *All the Myriad Ways* imagines a multiverse in which every individual's decision creates a fork of two parallel worlds, resulting in an exponential number of universes with vastly different levels of technology: "The Crosstime Corporation already held a score of patents on the inventions imported from alternate time tracks. Already those inventions had started more than one industrial revolution" (Niven 2001, p. 101). From a broad economic perspective, the innovation borrowed from parallel worlds could be positive. What if certain technologies replaced entire industries to good effect?

But the impact on individuals was different. For a pilot who traveled across alternate realities, for example, his "own world continued to divide after his departure, in a constant stream of decisions being made both ways" (Niven 2001, p. 104). Experiencing such a phenomenon led to suicides as people realized that if "alternate universes are a reality, then cause and effect are an illusion. The law of averages is a fraud. You can do anything, and one of you will, or did" (Niven 2001, p. 106).

Such a breakdown would have consequences as people did not see risk-reward and cause-effect in the same rational light. Economic theories tend to assume people act rationally on average, but a revelation about reality—even if untrue—can cause mass changes in behavior, breaking down theoretical models.

The Sterling-Shiner story *Mozart in Mirrorshades* (a reference to Mozart adopting the fashion of the future) depicts a future civilization of people who travel to the past of an alternate timeline to acquire natural resources, artifacts, and art. A fictionalized Thomas Jefferson says to the time travelers, "You made certain promises when we joined forces. You guaranteed us liberty and equality and the freedom to pursue our own happiness. Instead we find your machinery on all sides, your cheap manufactured goods seducing the people of our great country, our minerals and works of art disappearing into your fortresses, never to reappear!" (Sterling 2001, p. 300). This accelerated transition of technology in the past shows how some of the most-valuable goods were already present in the past. A slow transition, as done with history, masks some of the negative consequences and potential externalities of economic expansion.

Trade between the past and future has always occurred, in a sense: for instance, those who own works of art sell them for currency that can be used to purchase modern goods. Similar trade between cultures has also occurred, with results such as those in the story. A conventional wisdom is that as China has flooded much of the world with cheap manufactured goods, its wealthier citizens have used the proceeds to purchase luxury real estate all over the world.

In his story *Eutopia*, Poul Anderson explores travel to parallel universes. The protagonist compares a parallel society to his own: "They kept the population within bounds in Westfall as in Eutopia. But not because they knew that men need space and clear air, Iason thought. No, they acted from greed … A father did not wish to divide his possessions among many children" (Anderson 2001, pp. 257–258). Overpopulation is an issue explored in several aforementioned stories. In this case, Iason observes a natural correction to it, thanks to the greed (or, more generally, self-interest) innate to humans. While modern societies allow people to dictate how their assets are distributed after death, there have been nations where, for example, the law passes all possessions to the first-born son. In fictional Westfall, possessions are evenly divided among the heirs. This system of inheritance influences how individuals act.

Anderson makes an important point: “Society must have structure and meaning. But nature does not dictate what structure or what meaning” (Anderson 2001, p. 266). Many of society’s rules and norms are arbitrary, but adherence to such convention as exists is not arbitrary; for instance, while driving on the right side of the road is an arbitrary decision of government, the need to standardize a side is not.

However, sometimes choosing a particular side does make sense. If a race of people were physically asymmetrical, as are the aliens in *The Mote in God’s Eye*, one side of the road could very well be safer than the other, based on their anatomy.

Just because some custom or law has a purpose does not mean it is the ideal one. A monetary system is clearly helpful for an economy, as it allows transactions to occur more easily. But which monetary system works best? Commodity-backed currencies have given way to fiat money. Some argue “precious-metals”-backed currencies are superior because they do not allow inflation; but what happens when a gold-rich asteroid comes close enough to Earth that humans can mine it? What if a cheap technology can act as the philosopher’s stone, transforming common lead into gold? In that case, the gold-backed currency could become extremely devalued due to an oversupply of precious metals.

Studying different civilizations—foreign, historical, and yes, even fictional—is important. Certain arbitrary conventions could be impacting the economy in ways that were never apparent in the past but are transparent in science fiction.

24.7.4 History Repeated in the Future

Because most authors write in response to the world around them, the events and societies they depict often mirror those in which they live. Thus, futuristic science fiction depicts aspects of history repeating in the future: *A Canticle for Leibowitz* was based on the bombing of a monastery during World War II; *The Forever War* inspired by the Vietnam War; *District 9* an extension of District Six during South Africa’s period of apartheid.

Other works have predicted history *before* it happened. In *The Gods Themselves*, Isaac Asimov postulates a technology that takes advantage of differences in physics between two parallel worlds, generating seemingly free energy. But when scientists learn the transfer between universes could result in a collapse of the Sun, a Senator says, “Don’t ask me to stop the

Pumping. The economy and the comfort of the entire planet depend on it. Tell me, instead, how to keep the Pumping from exploding the Sun" (Asimov 1972, p. 56).

Modern society has become dependent on burning fossil fuels for electricity, while resources are both limited and a cause of climate change. The international goal of drastically reducing carbon emissions is not realistic with current technology. Electrifying all cars—a stated goal of various countries—would require an abundance of rare-earth metals that would create their own pollution. We suggest that no single action will solve the climate crisis. The solution will result from some combination of reducing emissions, shifting energy sources, capturing carbon from the air, and other changes—including some that have not yet been imagined outside of science fiction!

Asimov could not have seen today's climate crisis or energy squeeze: *The Gods Themselves* was published in 1972, just before the gas shortages. But Asimov did understand the concept of "no free lunch": there are potential externalities in many transactions.

After his death, Asimov's estate authorized other authors to publish a Second Foundation trilogy. In *Foundation's Triumph*, David Brin presents a philosophical discussion between two of Asimov's most-famous characters, the psychohistorian Hari Seldon and the robot R. Daneel Olivaw: "A certain faction of humans will always seek power over others. … We inherit this trait because those creatures who succeeded often had more descendants. … In ancient China, a powerful emperor could be relied on to check noble excesses. … The peaks and lows of aristocratic families made gaudy headlines, diverting the galaxy's masses, but … practical governance was left in the hands of meritocrats and civil servants. In psychohistorical terms, this was called an attractor state. In other words, society had a natural sink into which the power-hungry were drawn, fostering their preening illusions without doing much real harm. It had worked well for a long time in the Galactic Empire, much as it did in pretechnical China" (Brin 2000, p. 88).

Too many socioeconomic systems fail because they expect humans to act against their own nature. Brin recognizes a potential evolutionary trait that might have led to survival in ancient times but could be detrimental to a modern society. A stable civilization would depend on a political system in which those seeking power do not alter the operations affecting the broader society.

The conversation continues: “The Chinese created a special class of authorities who could only be loyal to the empire, and not to their own descendants. Because they would never have children … And an analogy in the modern Galactic Empire was obvious … Positronic robots programmed to think only of humanity’s good” (Brin 2000, p. 90). While modern society is not about to bring back eunuchs, and robots are far from advanced enough to play any administrative role, individuals’ motivation to serve their children and descendants could be curtailed through laws and other incentives. In general, as already noted, officials’ self-centered intentions are often at odds with their fiduciary duty to their constituents.

Once again, these individual motivations are potentially evolutionary. Humans who provided for their descendants’ survival were more likely to have their genes carried forward by those descendants. Even if this logic is false, a political and economic system that rewards individuals directly for beneficial contributions is likely to thrive and serve the wider population. A system that limits the concentration of power does not utilize resources, particularly human capital, to maximum efficiency. The equity-efficiency trade-off again!

Ensuring that wealth and power do not become more concentrated in the future is a worthy goal. Most countries have laws that tax income and inheritance, while encouraging spending in positive ways through philanthropy. Of course, maintaining a balance that rewards innovation against long-term power—or, more generally, the balance between equity and efficiency—is difficult. Reality is complicated; no model can fully capture it—nor can an anecdote via a story. Yet finding that balance is essential for the viability of a society. And studying both actual and imagined reality is a useful tool to do so.

24.8 Why Science Fiction Matters

From the innovative technologies that would allow Isaac Asimov’s *Foundation* to happen in the distant future, to the repeat of history examined in David Brin’s sequel that continued Asimov’s legacy, some component of economics plays a role in nearly every speculative story. In the end, psychohistory does not exist; there is not some statistical model that can predict any trend with near-certainty. Economics, too, does not provide some unified model that policymakers or businesses can depend on to make decisions affecting broad swaths of the population. Theories are just that: hypotheses subject to assumptions and initial conditions. Without adjusting those assumptions or understanding initial conditions, no theory is fit to be put into practice.

Science fiction, then, provides an abundance of thought experiments that twist the assumptions inherent in economic theory and practice. Economics uses models that approximate reality; but that reality is constantly shifting, and human nature has not been fully captured by any model. With history limited to a finite set of data, fiction expands on the tools available to academics and practitioners. Science fiction deals in potentially infinite alternate realities, ranging from natural extensions of human society—counterfactual history and tales of the near future—to those in the distant future across the galaxy.

From future changes in technology to the reimagining of historical events, theoretical tweaks to society lead to economic consequences. Available data are limited, and no experimental economics can be truly independent of the status quo. With human nature not fully understood, individuals' actions can vary in unpredictable ways. Authors may not comprehend human nature better than sociologists, but they can find exceptions to average behavior (unique actions make for good storytelling). Ultimately, economies and societies are not determined by averages: they are driven at the margins.

In today's politicized environment, where ideology can override common sense, science fiction is more important than ever. Even under ideal circumstances with best intentions, policies affect behavior in ways not easily predicted by data. We suggest that policymakers study science fiction to reduce inefficiencies and minimize disruption caused by their policies.

Science fiction can make policy more realistic because it imagines what can go wrong. Assuming the best-case scenario does not lead to best-practice legislation.

Like any thought experiment, a fictional story can never prove a theory true. However, it is an extremely useful tool to assess possible futures and their implications on decisions today. By becoming more versed in potential unintended consequences, economists working on policy and business can craft decisions the better to achieve objectives. While speculative fiction may provide many incorrect predictions of the future of humanity, its study could help ensure society's future prosperity.

Notes

1. We are not sharply distinguishing among "alternate history," "counterfactual history," and "speculative history;" nor are we exploring the development and methodologies of these genres. There is a substantial literature, among which we suggest Bunzl (2004), Turtledove (2001), McCloskey (2020), and Wikipedia (2021a).
2. Related works, not with our focus, are Evans (2013), Davies (2018), and Westfahl et al. (2020).
3. Officer (2019, 2002).
4. Needless to state, we are abstracting from the costs borne by the indigenous population in the Americas and by the enslaved persons brought from Africa therein. The conventional narrative (including here) of colonization of the Americas and American "manifest destiny" does not take these costs into account, but they are opportunity costs nevertheless.
5. The analogy to the mistreatment of the indigenous and enslaved people of the Americas is obvious.
6. Wikipedia (2021b) and data.worldbank.org.
7. https://fred.stlouisfed.org/series/LNS11300002.
8. Wikipedia (2021c), Tabuchi (2022). In all these cases, other geopolitical factors were also involved.
9. Vandenbroucke (2020).
10. This passage is a rudimentary analysis of inflation, for lack of space and for a straightforward correspondence with the the science-fiction coverage—a comment that applies to the other economic concepts in the chapter.
11. To say nothing of the dehumanization of enslaved persons.

References

Aldiss, Brian W. (1969). *Starship*. New York: Avon.
Amis, Kingsley (1976). *The Alteration*. New York: Viking.
Anderson, Poul (2001). "Eutopia." In Harry Turtledove, ed., *The Best Alternate History Stories of the 20th Century*, pp. 251–267. New York: Ballantine.
Asimov, Isaac (1972). *The Gods Themselves*. New York: Fawcett Crest.
——— (1985). *The Caves of Steel*. New York: Ballantine.
——— (2004). *Foundation*. New York: Bantam Dell.
Blomkamp, Neil (2009), director. *District 9* [Film]. QED International.
Bong, Joon-ho (2013), director. *Snowpiercer* [Film]. Moho Film.
Brin, David (2000). *Foundation's Triumph*. New York: HarperTorch.
Bunzl, Martin (2004). "Counterfactual History: A User's Guide." *American Historical Review* 109 (June), pp. 845–858.
Card, Orson Scott (2002). *Ender's Game*. New York: Tom Doherty.
Davies, William (2018), ed., *Economic Science Fiction*. London: Goldsmiths.

Dick, Philip K. (1982). *Blade Runner (Do Androids Dream of Electric Sheep?)*. New York: Ballantine.

——— (1988). *The Man in the High Castle*. New York: Ace.

Evans, Richard J. (2013). *Altered Pasts*. Waltham, MA: Brandeis University Press.

Farmer, Philip José (1972). "Riders of the Purple Wage." In Harlan Ellison, ed., *Dangerous Visions*, pp. 65–131. New York: Berkley Medallion.

——— (1980). *To Your Scattered Bodies Go*. Boston: Gregg.

Fleischer, Richard (1973), director. *Soylent Green* [Film]. Metro-Goldwyn-Mayer.

Gibson, William (1984). *Neuromancer*. New York: Ace.

Haldeman, Joe (2014). *The Forever War*. New York: Open Road Integrated Media.

Herbert, Frank (1990). *Dune*. New York: Ace.

Heinlein, Robert A. (1987). *Stranger in a Strange Land*. New York: Ace.

——— (2018). *The Moon Is a Harsh Mistress*. New York: Ace.

Huxley, Aldous (1998). *Brave New World*. New York: Perennial.

James, Phyllis D. (2006). *The Children of Men*. New York: Vintage.

Jemisin, Nora K. (2015). *The Fifth Season*. New York: Orbit.

Johnson, Rian (2012), director. *Looper* [Film]. TriStar Pictures.

Le Guin, Ursula K. (2000). *The Left Hand of Darkness*. New York: Ace.

——— (2006). *The Dispossessed*. New York: HarperCollins.

McCaffrey, Anne (1996). *Dragonflight*. New York: Del Rey.

McCloskey, Donald N. (2020). "Counterfactuals." In *The New Palgrave Dictionary of Economics*, 3rd ed. London: Palgrave Macmillan.

Miller, George (2015), director. *Mad Max: Fury Road* [Film]. Kennedy Miller Mitchell.

Miller, Walter M. (1961). *A Canticle for Leibowitz*. New York: Bantam.

Moore, Ronald D. (2019, 2021), producer. *For All Mankind* [Television series]. Tall Ship Productions.

Moore, Ward (2001). "Bring the Jubilee." In Harry Turtledove, ed., *The Best Alternate History Stories of the 20th Century*, pp. 152–249. New York: Ballantine.

Morgan, Richard K. (2006). *Altered Carbon*. New York: Del Rey.

Niccol, Andrew (2011), director. *In Time* [Film]. New Regency.

Niven, Larry (2001). "All the Myriad Ways." In Harry Turtledove, ed., *The Best Alternate History Stories of the 20th Century*, pp. 99–107. New York: Ballantine.

———, and Jerry Pournelle (1974). *The Mote in God's Eye*. New York: Simon & Schuster.

Nolan, Christopher (2014), director. *Interstellar* [Film]. Paramount Pictures.

Officer, Lawrence H. (2002). "The Bible Meets Science Fiction: A Solution to the Principal-Agent Problem." *Journal of Political Economy* 110 (February), back cover.

______. (2019). "Ghost Money in a Barter Economy." *Journal of Political* Economy 127 (June), back cover.

Orwell, George (1981). *Nineteen Eighty-Four*. Middlesex: Penguin.
Pohl, Frederik (1999). *Gateway*. London: Orion.
Rand, Ayn (1957). *Atlas Shrugged*. New York: Signet.
Robinson, Kim Stanley (2002). *The Years of Rice and Salt*. New York: Bantam.
Sheckley, Robert (1966). *Mindswap*. New York: Dell.
Silverberg, Robert (1981). *Lord Valentine's Castle*. New York: Bantam.
Simmons, Dan (1991). *Hyperion*. London: Headline.
Spottiswoode, Roger (2000), director. *The 6th Day* [Film]. Phoenix Pictures.
Sterling, Bruce, and Lewis Shiner (2001). "Mozart in Mirrorshades." In Harry Turtledove, ed., *The Best Alternate History Stories of the 20th Century*, pp. 296–312. New York: Ballantine.
Stephenson, Neal (2003). *Snow Crash*. New York: Bantam.
Tabuchi, Hiroko (2022). "Before War, Race Was on for Ukraine's Lithium." *New York Times*, March 8, p. D3.
Turtledove, Harry (1993). *The Guns of the South*. New York: Ballantine.
——— (2001). "Introduction." In Harry Turtledove and Martin H. Greenberg, eds., *The Best Alternate History Stories of the 20th Century*, pp. vii–xiii. New York: Ballantine.
Vinge, Vernor (2000). *A Deepness in the Sky*. New York: Tom Doherty.
Vandenbroucke, Guillaume (2020). *Which War Saw the Highest Defense Spending? Depends How It's Measured*, stlouisfed.org.
Vonnegut, Kurt (1970). *Player Piano*. New York: Avon.
——— (2006). *Cat's Cradle*. New York: Random House.
Wachowski, Lana, and Lilly Wachowski (1999), directors. *The Matrix* [Film]. Warner Bros.
Westfahl, Gary, Gregory Benford, Howard V. Hendrix, and Jonathan Alexander (2020). *Science Fiction and the Dismal Science*. Jefferson, NC: McFarland.
Wikipedia (2021a). *Alternate History*.
——— (2021b). *Digital Economy*.
——— (2021c). *Nickel Deposits of Finland*.
Wells, H. G. (1996). *The Time Machine*. London: Orion House.
Zelazny, Roger (1966). *This Immortal*. New York: Ace.
——— (2000). *Lord of Light*. New York: HarperCollins Avon Eos.

Ari J. Officer is a commodities trader for Jump Trading in Chicago. For some years he wrote columns for *TIME Magazine*. More recently, he authored several works of fantasy and science fiction: *Speculation* appears in the science-fiction anthology *Brave New Worlds*, edited by S. C. Butler and Joshua Palmatier, published by Zombies Need Brains, LLC (2022). *Reverse the Clock* was a finalist for the 2021 Baen Fantasy Adventure Award. *Music from the Stars* appears on Night Shift Radio's *The Storyteller Series* (2022).

Index of Names

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8

J

K

L

M

T

U

V

W

Y

Z

Index of Subjects

L. H. Officer, *Essays in Economic History*,
https://doi.org/10.1007/978-3-030-95925-8