Studies in Economic History

Jari Eloranta
Eric Golson
Andrei Markevich
Nikolaus Wolf *Editors*

Economic History of Warfare and State Formation

Springer

Studies in Economic History

Aim and Scopes

This series from Springer provides a platform for works in economic history that truly integrate economics and history. Books on a wide range of related topics are welcomed and encouraged, including those in macro-economic history, financial history, labor history, industrial history, agricultural history, the history of institutions and organizations, spatial economic history, law and economic history, political economic history, historical demography, and environmental history.

Economic history studies have greatly developed over the past several decades through application of economics and econometrics. Particularly in recent years, a variety of new economic theories and sophisticated econometric techniques—including game theory, spatial economics, and generalized method of moment (GMM)—have been introduced for the great benefit of economic historians and the research community.

At the same time, a good economic history study should contribute more than just an application of economics and econometrics to past data. It raises novel research questions, proposes a new view of history, and/or provides rich documentation. This series is intended to integrate data analysis, close examination of archival works, and application of theoretical frameworks to offer new insights and even provide opportunities to rethink theories.

The purview of this new Springer series is truly global, encompassing all nations and areas of the world as well as all eras from ancient times to the present. The editorial board, who are internationally renowned leaders among economic historians, carefully evaluate and judge each manuscript, referring to reports from expert reviewers. The series publishes contributions by university professors and others well established in the academic community, as well as work deemed to be of equivalent merit.

More information about this series at http://www.springer.com/series/13279

Jari Eloranta • Eric Golson • Andrei Markevich
Nikolaus Wolf
Editors

Economic History of Warfare and State Formation

Springer

Editors
Jari Eloranta
Department of History
Appalachian State University
Boone, NC, USA

Department of History
University of Jyvaskyla
Finland

Eric Golson
University of Surrey
Guildford, Surrey, UK

School of Oriental and African Studies
University of London
London, UK

Andrei Markevich
New Economic School
Moscow, Russia

Nikolaus Wolf
Humboldt University
Berlin, Germany

ISSN 2364-1797 ISSN 2364-1800 (electronic)
Studies in Economic History
ISBN 978-981-10-1604-2 ISBN 978-981-10-1605-9 (eBook)
DOI 10.1007/978-981-10-1605-9

Library of Congress Control Number: 2016953181

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Foreword

Roughly a generation after the fall of the Iron Curtain, our world seems to be less stable and predictable than ever. Many contemporary observers saw the collapse of the Soviet Union and the end of the Cold War as a sign of the ultimate victory of "Western" liberal democracy and free markets. And indeed, both democracy and market economies have gained ground after 1989, with democratic revolutions in many Eastern European countries; the extension and deepening of the European Union; a remarkable rise of China, India, and other Asian economies; or the revolutions of the Arab Spring. Not at least, sub-Saharan Africa has grown twice as fast as the world economy between 2005 and 2015. In terms of technology, the world has never been better connected, and markets have never been more deeply integrated than today. Many value chains are truly global today, when European customers buy US mobile phones, assembled in China using African raw materials. But while markets continue to expand and integrate, the institutional framework of our world seems to be increasingly incongruous and out of sync with economic developments. National policymakers are struggling to react to crises let alone to act proactively, and liberal democracies face challenges from growing inequality and aggressive ideologies.

What is going on? According to Dani Rodrik, modern societies face a choice between deep economic integration, national sovereignty, and democracy. You cannot have more than two of them together. So today, with unseen levels of global integration, the democratic nation–state is under stress, while elsewhere (large) authoritarian regimes seem to catch up. On an even more general level, this is nothing new. The economic history of the modern world, with a series of economic and political revolutions in Europe and her colonies since the eighteenth century and the ensuing Great Divergence, is a reminder: deep economic change provokes political conflict and institutional change and vice versa, not as an exception but as a rule. To understand how economic and institutional change interact, we need to analyze these periods of conflict, revolution, war, and state formation and to distill patterns and regularities out of often chaotic circumstances and incomplete (or outright manipulated) data. This is a huge challenge. The economist typically starts with a theoretical argument and more often than not prefers empirical work on periods of institutional stability to prove the point – an approach that is clearly in danger of circularity. The historian instead tends to look for the specific and to focus on contingencies. A problem here is the lack of a clear benchmark against which observations can be interpreted (to the extent that some historians have given up on the idea of objective knowledge altogether). What is needed is a combination of both, economic theory and historical methods, and someone with a mind open enough to find theories wrong and stubborn enough to spend weeks in archives looking for better data. Mark Harrison fits this description like few others. When asked what economic history means to him, he once paraphrased David Purdy that "instead of criticizing history in the light of our ideals, the thing is criticize our ideals in the light of history." With this program, he started in the late 1970s to explore the economics of the Soviet Union, the logic of coercion and its failure, and the economic history of World War I and II. The economist Harrison shows in "Coercion, Compliance, and the Collapse of the Soviet Command Economy," published in the *Economic History Review* in 2002, how microeconomic theory can help to shed new light on the stability and ultimate collapse of the Soviet command economy. In a paper with Andrei Markevich (*Journal of Economic History* 2011), the historian Harrison produced new estimates of Russia's national income during the chaotic period from 1913 to 1928, based on a critical analysis of new data. In his recent work on the KGB, Mark Harrison combines both original archival work on the ultimate type of difficult data – classified KGB files – and economic methods to assess and benchmark the costs of secrecy ("Accounting for Secrets," *Journal of Economic History* 2013).

The chapters in the present book are written by scholars who have all been inspired by Mark's work, as colleagues or students. The various contributions in the *Economic History of Warfare and State Formation* analyze issues that we need to understand better if we want to deal with current and future challenges. How did states and societies survive crisis and war in the past? How will economic development shape institutional change? What type of institutions foster development and participation and under what conditions will these institutions dominate others? All these questions sound very important and heavy, but understanding patterns from history can be fascinating and indeed great fun. "... When I started doing economic history I thought that history was all in the past; I didn't understand that it was still going on" (Mark Harrison).

Humboldt University
Berlin, Germany

Nikolaus Wolf

Preface

The theoretical and methodological impact of economic history in the study of war, state formation, and the development regime types has been profound, especially in the last 40–50 years. Once seen as the domain of political scientists, economic theory, historical analysis, and quantitative methods are now used to explain the choices and outcomes of war. Unfortunately, these studies have often been scattered in various journals, and the topics and authors have not necessarily interacted with each other very effectively. Political scientists and economists may not always be aware of the contributions made by economic historians in these profoundly important fields and topics. However, they have not always been interested in the interdisciplinary aspects which economic historians bring to the field. For example, most studies in defense economics have been limited in analytical time span, focusing on 1945 onward. Longer-run historical issues have typically fallen outside the interest of defense economists, although many of the same tools and theoretical insights are useful for such long-run analysis. Political and conflict scientists, including peace sciences, often cover similar ground as defense economics but focus more on the causal factors behind the most destructive conflicts and the determinants of state formation.

In this volume, we are bringing together a diverse group of economic historians, economists, and political scientists to shed light on some of these issues. Despite the different educational and institutional backgrounds of the authors, we all have in common a desire to explain big issues in history using interdisciplinary methods. The types of big issues include the formation of European and non-European states in the early modern period, the emergence of fiscal states and eventually modern democracies with extensive welfare states, the violent upheavals that influenced these processes, the persistence of dictatorships and nondemocratic forms of government, and the arrival of total war and its consequences. They demonstrate the long-run dynamics of military spending and warfare. As Charles Tilly (1990) suggests, one of the key contributors in the study of state formation, *coercion* (a monopoly of violence by rulers and an ability to also wield coercion externally) and *capital* (the means of financing warfare) were the key elements in the European ascendancy to world domination in the early modern era. Warfare, state formation, and technological supremacy were all interrelated fundamentals of the same process.

Warfare and State Formation

One of the motivations of this volume is to bring forth interdisciplinary scholarship on warfare, conflicts, and state formation, especially to reflect on the research undertaken by Mark Harrison in his career. He has made revolutionary contributions to the topic of warfare and state formation both in terms of cross-country comparisons and country-specific case studies. His research on the two world wars has extended our understanding of economics and economic history of these global conflicts. He has inspired and edited two volumes on two world wars, each of which put the experiences of individual countries into international perspective (the volume on World War I was co-edited with Stephen Broadberry; see Harrison 1998a; Broadberry and Harrison 2005). Based on these volumes, Harrison (1998b) has compared and summarized national stories underlining economic factors behind military success in these global wars (see also Harrison 2015). He demonstrated it was more likely for wealthy nations to win these conflicts because of the larger share of national income available to them for mobilization for warfare and that geography mattered in the sense it could provide time for such transformation while other factors were largely conditional on wealth, geography, and time. Harrison's (1985, 1996, 2008; Barber and Harrison 1991) country-specific case studies demonstrated there is an exception to these rules, which worked only for market economies. National governments tended to employ elements of strategic planning and command management during global wars.

Following on this, Mark Harrison was one of the first scholars who took advantage of the opening of former Soviet archives to investigate the political economy of the command system and Stalin's dictatorship, including the effect of the warfare and military threats on consolidation of power in interwar Soviet Union. He reconstructed national accounts of Russia and the USSR in both world wars (Harrison 1996; Markevich and Harrison 2011) and found the Soviet economy was very close to collapse in 1942; it managed to recover because the command system had significant comparative advantages in mobilizing recourses for battles. In contrast, the Russian economy in World War I fits international patterns. In the interwar period, international politics and potential military threats shaped the speed of Soviet industrialization. The government directed recourses toward heavy industry and the defense sector, thereby increasing military expenditures at the cost of other sectors and consumption (Davies, Harrison and Wheatcroft 1993; Davies and Harrison 1997; Gregory and Harrison 2005). However, despite its priority status, the defense sector suffered from all the main drawbacks of the command model: agents behaved opportunistically, they manipulated information, and they colluded (Harrison 1998c, 2003, 2005; Barber and Harrison 2000; Markevich and Harrison 2006). While accumulating tremendous power, the dictator could not solve the fundamental problem of command to secure implementation of all his orders (Harrison 2005). Harrison (2008) argues that, in the Soviet context of the 1930s, repression and military capacity could substitute rather than complement each other. Stalin feared that external and internal enemies would cooperate in future open conflict against him. To prevent such conflict and potential union, the dictator could decrease the chances of his potential enemies by increasing the strength of his own regime either by eliminating potential dissidents or building military power. Political repressions were easier to realize in the short run, while the second option required time. Indeed, the dynamic of Stalin's repressions illustrated that the number of political victims grew in response to deterioration of international relations and the increase of the likelihood of war in the early and late 1930s, being substituted by development of the defense sector some time later after foreign shocks (Harrison 2008). The threat of coercion and repression played a key role in the stability of the regime, and once it became non-credible, the regime failed (Harrison 2002). In a similar vein, Andrei Markevich in his chapter in this volume explores patterns of repressions and punishment under Stalin using micro-evidence extracted from the Soviet archives.

In general, warfare has affected state formation in dictatorships as well as in democracies. There are many examples in world history when rising autocrats used external threats as a tool to get internal political support and to accumulate power. Similar, dictators often employed political mobilization due to incoming imaginary or real military risks to maintain the regime's stability that might be a cheaper option than two other main tools at their hands, repressions and purchases of loyalty (Wintrobe 1998). Simultaneously, such policy made hostile collective actions more expensive (Tullock 2005). In other words, warfare could contribute to the formation of vested interests and extractive institutions, rather than inclusive institutions. The effect of military conflicts on development is ambiguous since they could prevent the formation of wide political coalition and broad distribution of power in the society, a critical condition for the emergence of inclusive political institutions and subsequent economic success (Acemoğlu and Robinson 2012 and their chapter in this volume).

Russian and Soviet economic history has been long one of Mark Harrison's professional interests, and he has contributed significantly to the field. Originally, the nomadic challenge from the East led to the rise of the autocratic Moscow state in the middle of the second millennium and to the concentration of the bulk of available recourses under the control of the central government. In particular, property rights on land were limited and conditional to military service of the owner of a land title which had long-run consequences for both the distribution of power and market development (Pipes 1974; Owen 1995). Similar, the emergence of the modern state under Peter the Great in the early eighteenth century was driven by military conflict with Sweden and the necessity to launch a state capable of taxing the population efficiently to support a modern army. However, in contrast to its European counterparts working on similar tasks, Russia managed to build a fiscal system that relied more on bureaucracy and coercion rather than on taxation of market transactions (Tilly 1990). As Osinsky and Eloranta argue in their chapter in this volume, the military factors were an important reason which caused a divergence between Russia and Finland after the Russian revolution of 1917. The ability of the Bolsheviks to attract rank-and-file soldiers of the former imperial army to their side and the failure of Red Finns in this same respect were crucial to the formation of governing coalitions in the two countries. Finally, military considerations, and in particular the war alarm of 1927, were used by Stalin in the struggle with his political opponents as a crucial argument in favor of rapid industrialization and the creation of the command economy (Simonov 1996) and in this way contributed to the emergence of his personal dictatorship (Harrison 2008).

The Russian government initiated reforms to transform institutions and to expand the distribution of power in response to military failures; in contrast, successful wars blocked changes of dominantly extractive institutions. Alexander II launched the "Great Reforms" of the 1860s and 1870s after the defeat in the Crimean War when the relative economic backwardness of the country became obvious for the government and the public (Gerschenkron 1962). The institution of serfdom was abolished in 1861 despite the opposition of the gentry, which benefited from this system extracting rents from the peasantry at the costs of slower national development both in the middle and long runs (Buggle and Nafziger 2016; Markevich and Zhuravskaya 2016). The "Great Reforms" established self-government institutions at the regional and local levels in the empire that extended state capacity as Nafziger shows in his chapter in this volume. Similar, the lack of clear success in the Soviet invasion of Afghanistan and the Cold War contributed to Gorbachev's decision to adjust the command economy and to start the *Perestroika* that eventually triggered the transition process. In other cases, the victories over Napoleon or Nazi Germany stabilized existing autocratic regimes.

There are broader lessons in the economic history of state formation, too, for example, from the Asian cases. As Gupta et al. show in this volume, the Asian states prior to European colonial exploits were grappling with the themes of state formation and the expansion of military capacity. However, as they point out, state capacity may be linked to the emergence of wide differences in economic and military power. Philip Hoffman has recently also explored this dimension and argued that the Asian states did not persistently experience constant warfare (i.e., competition), expand and maintain their revenue base, and invest in military capacity (which, in turn, would lead to technological innovations in the defense sector). The European experience of a near-constant warfare (=tournament) with high "prizes" to be captured is one of the keys to unravel why the Great Divergence emerged (Hoffman, 2015; see also Broadberry and Gupta 2006; Pomeranz 2000, on the Great Divergence debate). Gupta et al. in their chapter, of course, recommend caution in making such broad comparative claims.

Total War and Conflicts

Another key theme in this volume concerns the economic history of the world wars. Specifically, in this volume, we are focused on analyzing the issue of economic and military mobilization more closely. In essence, we want to examine one of the key components in world history, i.e., the concept of total war. As argued by David Bell (2007), the French Revolutionary Wars and the ensuing Napoleonic conflicts should be placed into the same category as the world wars, given the global nature of the conflicts, the use of modern weaponry, the organization of war economies and occupations, and the civilian casualties. World War I, which began in August 1914, was an even more massive conflict, aided by expanded military–industrial capacity in the West especially. New studies have focused on the mobilization during the war and the impact of the war on the civilian populations (see, e.g., Blum 2011, 2013). Similarly, there has been ample scholarship on the economic dimensions of World War II (in addition to early references, see especially Budrass et al. 2010; Scherner 2010; Tooze 2007).

So, what happened? Why did the twentieth century become so deadly and why did we see the world wars at this time? According to Charles Tilly, the reasons included the rapid development of more deadly weapons, along with more centralized and powerful nation–states (Tilly 1990). Niall Ferguson pointed out that the average yearly amount of (Great Power) war was highest in the sixteenth and lowest in the nineteenth century. During the nineteenth and twentieth centuries, in fact, the industrialized nation–states found new ways to mobilize their manpower and resources for warfare, and the technological advances of the age, in particular railroads, often served both the civilian as well as military production and planning. (Ferguson 1999, 2006). The world wars were unparalleled in their severity and concentration and pioneered the widespread use of genocides, or democides, in this century (Ferguson 2006; Rummel 1997). World War I (1914–1918) featured more than 30 countries fighting a protracted total war, leading to the death of 20 million people, only to be dwarfed by World War II (1939–1945), in which more than 60 countries waged war and prematurely ended the lives of more than 55 million people. The European region, as well as many other battlefields around the globe, was decimated by these conflicts, a situation that did not improve until after World War II.

The newer studies have now put more emphasis on analyzing the specific impacts of the world wars on particular countries and on various sections of the populace. Moreover, mobilization for the war prior and during the conflict and the impacts of demobilization have now received more attention in the literature (see, e.g., Hantke and Spoerer 2010; Ritschl 2004, 2005). This is an area that this volume will contribute to. In particular, here we offer an emphasis on, e.g., the neutral states, on specific war economies, and on comparing the experiences of countries in the two world wars (see Eloranta and Harrison 2010). While most of the chapters here focus on the short-run issues, Hugh Rockoff's chapter, for example, examines the US experience at war for several centuries. It seems that during smaller conflicts the United States was able to fund the war by borrowing and increased taxation, whereas during major conflicts a limit was reached and printing money (and thus inflation) became another tool for financing the war.

In terms of the chapters on the world wars, this volume features a variety of perspectives on these conflicts. For example, Mark Harrison's chapter provides a re-evaluation of the very nature of World War I. Contrary to a lot of the existing literature, he argues the onset of the war was not an accident: key players in the war cabinets knew very well this would be a massive and lengthy conflict, that the lethality of the war in the west was intentional, and that the Versailles Peace Treaty was not really a cause for the next world war. The rest of the chapters, in fact, focus on World War II. Eric Golson examines three neutral states (Spain, Sweden, and Switzerland), and he argues that all three had to adjust to the economic and political realities of the period, granting calculated concessions to both the Axis and the Allies. While they had to do that, particularly the Axis powers were quite dependent on the raw materials and resources the neutral states had to offer. Broadberry and Howlett offer lessons from the British mobilization experience during the war. The British decision-makers learned from mobilization in World War I and mobilized the economy to a greater extent and faster during World War II. Inflation was kept better in check and they relied on external assistance, borrowing more. On the whole, the war was naturally also much more costly and destructive for Britain. Fishback and Jaworski's chapter analyzes the US case during World War II, and the increased scale of the conflict is again emphasized by the authors. Furthermore, based on their model and empirical findings, they argue that the US mobilization was mostly influenced by changes in local economies and populations, due to fiscal expansion during the war. They also advocate for further studies of the mobilization at the various levels (federal, state, local) and the interactions with the political economy. Finally, Scherner and Streb examine the Axis side, by looking into the so-called German armament miracle supposedly engineered by Albert Speer. After carefully looking into other data and supplementing the figures, they find no change in the mobilization in 1942, when Speer became the armament minister, whereas armament production was adversely impacted by the beginning of the invasion of the Soviet Union in 1941, as resources were diverted to support the operation. Scherner and Streb's conclusions show politics and political economy are important in analyzing wartime mobilization.

We hope that this volume can inspire others to look into these issues and challenge and complement our work in this regard. Moreover, we hope other readers will be inspired to look into the scholarship of Mark Harrison to examine these complex issues. We have certainly benefited from the insights arising from his work and the personal inspiration he has given to many of us about our own work. We hope to be able to shed new light on the economic and historical dimensions of war and state formation, a topic that is still very much discussed and debated among scholars in different disciplines.

Boone, NC, USA — Jari Eloranta
Moscow, Russia — Andrei Markevich
London, UK — Eric Golson

References

Acemoğlu, Daron, and James A. Robinson. 2012. *Why nations fail*: *The origins of power*, *prosperity and poverty*. New York: Crown.

Barber, John, and Mark Harrison. 1991. *The Soviet home front*, *1941–1945*: *A social and economic history of the USSR in World War II*. London: Longman Publishing Group.

Barber, J., and M. Harrison. (eds.). 2000. *The Soviet defence-industry complex from Stalin to Khrushchev*. Basingstoke: Palgrave MacMillan.

Bell, David. 2007. *The First Total War*: *Napoleon*'*s Europe and the Birth of warfare as we know it*. New York: Houghton Mifflin Harcourt.

Blum, Matthias. 2011. Government decisions before and during the First World War and the living standards in Germany during a drastic natural experiment. *Explorations in Economic History* 48(4): 556–567.

Blum, Matthias. 2013. War, food rationing, and socioeconomic inequality in Germany during the First World War. *The Economic History Review* 66(4): 1063–1083.

Broadberry, Stephen, and Bishnu Gupta. 2006. The early modern great divergence: Wages, prices and economic development in Europe and Asia, 1500–1800. *Economic History Review* 59: 2–31.

Broadberry, Stephen, and Mark Harrison. 2005. The economics of World War I: An overview. In *The economics of World War I*, ed S. Broadberry and M. Harrison, Cambridge: The Cambridge University Press.

Budrass, Lutz, Jonas Scherner, and Jochen Streb. 2010. Fixed-price contracts, learning, and outsourcing: Explaining the continuous growth of output and labour productivity in the German aircraft industry during the Second World War. *The Economic History Review* 63(1): 107–136.

Buggle, Johannes, and Steven Nafziger. 2016. *Long-run consequences of labor coercion*: *Evidence from Russian serfdom*. Working paper.

Barber, John, and Mark Harrison. 1991. *The Soviet home front*, *1941–5*: *A social and economic history of the USSR in World War II*. London: Longman.

Barber, John, and Mark Harrison. (eds.). 2000. *The Soviet defence industry complex from Stalin to Khrushchev*. Basingstoke and London: Macmillan Press.

Davies, R.W., Mark Harrison and S.G. Wheatcroft. (eds.). 1993. *The economic transformation of the Soviet Union*, *1913–1945*. Cambridge: Cambridge University Press.

Davies, R.W., and Mark Harrison. 1997. The Soviet Military-Economic Effort under the Second Five-Year Plan (1933–1937). *Europe-Asia Studies* 49(3): 369–406.

Eloranta, Jari, and Mark Harrison. 2010. War and Disintegration, 1914–1950. In *The Cambridge economic history of modern Europe. Volume 2*: *1870 to the present*, ed. S. Broadberry and K. H. O'Rourke, Cambridge: Cambridge University Press.

Ferguson, Niall. 1999. *The pity of war. Explaining World War I*. New York: Basic Books.

Ferguson, Niall. 2006. *The war of the world*: *Twentieth-century conflict and the descent of the West*. London: Allen Lane.

Gerschenkron, Alexander. 1962. *Economic backwardness in historical perspective. A book of essays*, Cambridge, MA: Belknap Press of Harvard University Press.

Gregory, Paul, and Mark Harrison. 2005. Allocation under dictatorship: Research in Stalin's archives. *Journal of Economic Literature* 43(3): 721–761.

Hantke, Max, and Mark Spoerer. 2010. The imposed gift of Versailles: The fiscal effects of restricting the size of Germany's Armed Forces, 1924–9. *The Economic History Review* 63(4): 849–864.

Harrison, Mark. 1996. *Accounting for war*: *Soviet production*, *employment*, *and the defence burden*, *1940–1945*. Cambridge: Cambridge University Press.

Harrison, Mark. 1985. *Soviet planning in peace and war 1938–1945*. Cambridge: Cambridge University Press. Hardback edition.

Harrison, Mark. (ed.). 1998a. *The economics of World War II. Six great powers in international comparisons*. Cambridge,: Cambridge University Press.

Harrison, Mark. 1998b. The economics of World War II: An overview. In *The economics of World War II. Six great powers in international comparisons*, ed. M. Harrison. Cambridge: Cambridge University Press.

Harrison, Mark. 1998c. Prices, planners, and producers: an agency problem in Soviet industry, 1928–1950. *Journal of Economic History* 58(4): 1032–1062.

Harrison, Mark. 2002. Coercion, compliance, and the collapse of the Soviet command economy. *Economic History Review* 55(3): 397–433.

Harrison, Mark. 2003. The political economy of a Soviet military R&D failure: Steam power for aviation, 1932 to 1939. *Journal of Economic History* 63(1): 178–212.

Harrison, Mark. 2005. The fundamental problem of command: Plan and compliance in a partially centralised economy. *Comparative Economic Studies* 47(2): 296–314.

Harrison, Mark. (ed.). 2008. *Guns and rubles*: *The defense industry in the Stalinist State*. New Haven: Yale University Press. 2008.

Harrison, Mark. 2015. *The economics of coercion and conflict*. London: World Scientific Publishing.

Hoffman, Philip T. 2015. *Why did Europe conquer the world*? Princeton: Princeton University Press.

Markevich, Andrei, and Mark Harrison. 2006. Quality, experience, and monopoly: The Soviet market for weapons under Stalin. *Economic History Review* 59(1): 113–142.

Markevich, Andrei and Mark Harrison. 2011. Great war, civil war, and recovery: Russia's national income, 1913 to 1928. *Journal of Economic History* 71(3), 672–703.

Markevich, Andrei, and Ekaterina Zhuravskaya. 2016. *The economic effects of the abolition of serfdom. Evidence from the Russian empire*. SSRN Working paper: http://ssrn.com/abstract=2514964.

Owen, Thomas. 1995. *Russian corporate capitalism from peter the great to Perestroika*. New York: Oxford University Press.

Pipes, Richard. 1974. *Russia under the old regime*. New York: Charles Scribner's Sons.

Pomeranz, Kenneth. 2000. *The great divergence*: *Europe*, *China*, *and the making of the modern world economy*. Princeton: Princeton University Press.

Ritschl, Albrecht. 2004. The Marshall Plan, 1948–1951. *EH. Net Encyclopedia*.

Ritschl, Albrecht. 2005. The pity of peace: Germany's economy at war, 1914–1918 and beyond. In *The economics of World War I*, ed. S. Broadberry and M. Harrison, Cambridge: Cambridge University Press.

Rummel, R.J. 1997. *Power kills. Democracy as a method of nonviolence*. New Brunswick: Transaction Publishers.

Scherner, Jonas. 2010. Nazi Germany's preparation for war: Evidence from revised industrial investment series. *European Review of Economic History* 14(3): 433–468.

Simonov, Nikolaj. 1996. *Voenno-promyshlennyi kompleks SSSR v 1920–1950e gody*: *tempy ekonomicheskogo rosta*, *struktura*, *organizatsiia proizvodstva i upravlenie*. Moscow: Rosspen.

Tilly, Charles. 1990. *Coercion*, *capital*, *and European states*, *AD 990–1990*. Cambridge, MA: Basil Blackwell.

Tooze, Adam. 2007. *The wages of destruction*: *The making and breaking of the Nazi economy*. London: Penguin.

Tullock, Gordon. 2005. *The selected works of Gordon Tullock*, vol. 1–10, ed. Charles Rowley. Indianapolis: Liberty Fund.

Wintrobe, Ronald. 1998. *Political economy of dictatorship*. Cambridge and New York: Cambridge University Press.

Contents

Contributors

Editors

Jari Eloranta is a professor of comparative economic and business history at Appalachian State University in Boone, North Carolina, USA, as well as adjunct professor of history at the University of Jyvaskyla in Finland. He is also the secretary general of the International Economic History Association. He received his Ph.D. in economic history from the European University Institute in Florence, Italy, in 2002. He has published widely on the following topics: cost and impact of conflicts, history of military spending, development of nineteenth- and twentieth-century states, small states and trade in the last three centuries, Nordic economic history, and twentieth-century civil wars.

Eric Golson a researcher in the area of economic warfare, is currently a research fellow at Oxford University as well as a teaching fellow in economics at the University of Surrey and the School of Oriental and African Studies, London. He has previously taught at the London School of Economics (LSE) and the University of Warwick's Economics Department. In 2015, he earned his fifth teaching award, this time for undergraduate teaching at Warwick, and in 2013 became the first faculty member at the LSE to have received three awards for outstanding teaching. He received his Ph.D. from the LSE and both his bachelor's and master's degrees from the University of Chicago.

Andrei Markevich is an associate professor at the New Economic School in Moscow. Andrei specializes in economic history, with a particular emphasis on institutional aspects of Russian and Soviet economic development and political economy. His current project explores the interaction between institutional reforms, internal conflicts, and economic development, using examples from Russian history before and after the 1917 Revolution. He has published articles in the *Journal of Development Economics*, *Journal of Public Economics*, *Journal of Economic History*, and *Journal of Comparative Economics*. His paper on Russian national income in 1913–1928 was awarded the Russian National Prize in Economics in

2011. He received his Ph.D. from the Institute of Russian History, Russian Academy of Science, in 2002. He was a Marie Curie research fellow at the University of Warwick from 2005 to 2007.

Nikolaus Wolf is a professor of economics at Humboldt University Berlin, Germany, and research fellow at CEPR (London) and CESifo (Munich). He studied history and economics in Freiburg, Berlin, Gdansk, and London and held positions at the London School of Economics (UK), the Free University Berlin (Germany), and the University of Warwick (UK) before joining Humboldt in 2010. The focus of his work is on international economics in a long-run perspective, especially on trade, economic geography, and monetary policy. Recent publications include "The Economics of Density: Evidence from the Berlin Wall" (*Econometrica* 2015) and "The Location of the British Cotton Textiles Industry in 1838: A Quantitative Analysis" (*Journal of Economic History* 2014). Since 2015, he has been the editor-in-chief of the *European Review of Economic History*.

Authors

Daron Acemoğlu is the Elizabeth and James Killian professor of economics in the Department of Economics at the Massachusetts Institute of Technology. He received his Ph.D. in economics at the London School of Economics in 1992. He is an elected fellow of the National Academy of Sciences (USA) and several other similar organizations. He has received numerous awards and fellowships, including the John Bates Clark Medal in 2005. His areas of research include political economy, economic development and growth, human capital theory, growth theory, innovation, search theory, network economics, and learning. In addition to scholarly articles, Daron Acemoğlu has published four books: *Economic Origins of Dictatorship and Democracy* (jointly with James A. Robinson; *Introduction to Modern Economic Growth*; *Why Nations Fail*: *The Origins of Power*, *Prosperity and Poverty* (jointly with James A. Robinson); and *Principles of Economics* (jointly with David Laibson and John List).

Stephen Broadberry is a professor of economic history at Nuffield College, Oxford, research theme leader at CAGE, director of the economic history program at CEPR, and a Hans Christian Andersen professor at the University of Southern Denmark. He has also taught at the London School of Economics and at many other institutions. His research interests include the development of the world economy from AD 1000 to the present, sectoral aspects of comparative growth and productivity performance during the nineteenth and twentieth centuries, and wars and economic performance. He has been editor of the *Economic History Review* and also editor of the *European Review of Economic History*. He has been president of the European Historical Economics Society and has recently been elected president of the Economic History Society (from April 2016). His books include *The British*

Economy Between the Wars: *A Macroeconomic Survey* (Blackwell, 1986); *The Productivity Race*, *1850–1990*: *British Manufacturing in International Perspective*, *1850–1990* (CUP, 1997); *Market Services and the Productivity Race*, *1850–2000*: *British Performance in International Perspective* (CUP, 2006); the two-volume *Cambridge Economic History of Europe*, edited with Kevin O'Rourke (CUP, 2010); and *British Economic Growth*, *1270–1870*, co-authored with Bruce Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen (CUP, 2015).

Price V. Fishback is the Thomas R. Brown professor of economics at the University of Arizona and a research associate with the National Bureau of Economic Research. He is the current executive director of the economic history association and served as co-editor of the *Journal of Economic History* from 2008 to 2012. His co-authored books published by the University of Chicago Press include *Well Worth Saving*: *How the New Deal Safeguarded Homeownership* (2013), *Government and the American Economy*: *A New History* (2007), and *A Prelude to the Welfare State*: *The Origins of Workers*' *Compensation* (2000). The latter book was awarded a Samuelson Certificate of Excellence by the TIAA-CREF Institute and the Lester Prize for an Outstanding Book in Labor Economics and Industrial Relations by the Industrial Relation Sections at Princeton University. His first book was *Soft Coal*, *Hard Choices*: *The Economic Welfare of Bituminous Coal Miners*, *1890–1930* with Oxford University Press in 1992.

Bishnupriya Gupta is a professor in the Economics Department at the University of Warwick and a research associate of the Centre for Competitive Advantage in the Global Economy and a research fellow of the Centre of Policy Research. Her research has focused on industrialization in colonial India and the Great Divergence between Europe and Asia. Her recent publications include "India and the Great Divergence: An Anglo-Indian Comparison of GDP Per Capita, 1600–1871," with Stephen Broadberry and Johann Custodis in *Explorations in Economic History* (2015, 55:58–75), and "Discrimination or Social Networks? Industrial Investment in Colonial India" in the *Journal of Economic History*. She is one of the editors of *A New Economic History of Colonial India* (London: Routledge, 2015).

Mark Harrison is a professor of economics at the University of Warwick; research fellow at the Centre for Russian, European and Eurasian Studies at the University of Birmingham; and research fellow of the Hoover Institution on War, Revolution, and Peace at Stanford University. His research is on the economic history of Russia, communism, conflict, and security. He received the Alec Nove Prize for Russian, Soviet, and post-Soviet studies in 1997 and the Russian National Prize for Applied Economics in 2012. His latest book is *One Day We Will Live Without Fear*: *Everyday Lives Under the Soviet Police State* (Hoover Institution Press, 2016).

Peter Howlett is an associate professor in the Economic History Department at the London School of Economics. He received his Ph.D. from the University of Cambridge and before joining the LSE was a lecturer in economics at the University of Stirling. He has recently completed a stint as secretary of the Economic History

Society. He has published extensively on the British economy in World War II, including *Fighting with Figures* (with the Central Statistical Office) (HMSO, 1995), and has also written on World War I. Besides war economies, his other research interests include income distribution; growth, trade, and convergence in the international economy since 1870; internal labor markets (especially in the UK in the late nineteenth century); business history; and technology transfer in contemporary Indian agriculture. He was involved in a major research project on the dissemination of reliable knowledge, which led to the volume *How Well Do Facts Travel*? (CUP, 2011), co-edited with Mary Morgan.

Taylor Jaworski is an assistant professor of economics at Queen's University in Kingston, Ontario. His research focuses on World War II, regional development policy, and structural transformation in the USA. He received his Ph.D. in economics from the University of Arizona, holds an M.Sc. in economic history from the London School of Economics, and a B.A. in economics from George Washington University.

Debin Ma is an associate professor of the Department of Economic History at the London School of Economics, with research affiliation with the Australian National University, the University of Warwick, and Fudan University. He served as the secretary general of the International Economic History Association from 2012 to 2015. He obtained his Ph.D. in economics from the University of North Carolina at Chapel Hill and has taught in the USA, Japan, and China. His research interests include long-term economic growth in East Asia, international comparison of living standards, human capital and productivity, and comparative institutions, legal traditions, and long-run growth. He recently worked with Loren Brandt and Thomas G. Rawski and published an extended review essay titled "From Divergence to Convergence: Re-evaluating the History Behind China's Economic Boom" in the *Journal of Economic Literature*.

Steven Nafziger is an associate professor of economics at Williams College (USA) and a center associate at the Davis Center for Russian and Eurasian Studies at Harvard. He is an economic historian with broad and comparative research interests in how institutional factors help explain the pace and structure of imperial Russian economic development. Recent and ongoing projects have explored the economics of serfdom and emancipation, prostitution and female labor market opportunities, rural property rights, and the role that the structure and financing of corporations played in the early stages of Russian industrial development.

Pavel Osinsky is an associate professor of sociology at Appalachian State University in Boone, North Carolina, USA. Born and raised in Buryatia, one of Russia's autonomous republics in Eastern Siberia, Dr. Osinsky received his Ph.D. in sociology from Northwestern University (2007). His primary field of specialization is comparative–historical sociology, with a particular interest in examining the impact of the world wars of the twentieth century on the revolutionary cataclysms and the communist state-building in Eastern Europe and East Asia. His articles have been published in the *Annual Review of Sociology*, the *Sociological Forum*, the *Sociological Quarterly*, *Political Power and Social Theory*, *Research in Economic History*, and other academic journals.

James A. Robinson is a university professor at the Harris School of Public Policy at the University of Chicago. He was formerly the Wilbur A. Cowett professor of government at Harvard University. He studied economics at the London School of Economics, the University of Warwick, and Yale University. His main research interests are in comparative economic and political development with a particular interest in Latin America and sub-Saharan Africa. He is co-author with Daron Acemoğlu of the books *Economic Origins of Dictatorship and Democracy* and *Why Nations Fail*: *The Origins of Power*, *Prosperity and Poverty*. He is currently conducting research in the Democratic Republic of the Congo, Sierra Leone, Haiti, and Colombia, where he has taught for many years during the summer at the University of the Andes in Bogotá.

Hugh Rockoff is a distinguished professor of economics at Rutgers, the State University of New Jersey, and a research associate of the National Bureau of Economic Research. His research focuses on the history of monetary policy and financial regulation and the economics of war. He is the author of books on the US Free Banking Era and the history of price controls in the USA and of many papers in professional journals. His latest book is *America's Economic Way of War*: *War and the US Economy from the Spanish-American War to the Persian Gulf War*. He is also the author, with Gary Walton, of a textbook: *History of the American Economy*. His undergraduate degree is from Earlham College and his Ph.D. from the University of Chicago.

Tirthankar Roy is a professor of economic history at the London School of Economics. His research interests are modern South Asian history and development and global economic history. His recent publications include *Law and the Economy of Colonial India* (University of Chicago Press, 2016), *An Economic History of Early Modern India* (Routledge, 2013), and *India in the World Economy from Antiquity to the Present* (Cambridge University Press, 2012).

Jonas Scherner is a professor in modern European economic history at the Norwegian University of Science and Technology (NTNU) in Trondheim. He studied economics and history at the University of Mannheim and the University of Parma and received both his doctoral degree in economic history and his habilitation degree in economic and social history from the University of Mannheim. In the academic year 2005–2006, he was a visiting fellow at the Economic Growth Center, Yale University, and in 2008–2009, visiting research fellow in economic and social history at the German Historical Institute, Washington, DC. From 1995 until 2010, he worked in several positions at the University of Mannheim, before joining the NTNU in 2010, first as an associate professor and from 2011 on as professor. His current research focuses on the economic history of the Nazi period, especially the exploitation of occupied Europe during Nazi rule; on war economies, with a special focus on the German raw material policy; and on the reconstruction of historical data. His recent publications include “Armament in the Depth or Armament in the Breadth? German Investments Pattern and Rearmament During the Nazi Period”. (*Economic History Review* 66(2) 2013).

Jochen Streb is a professor of economic history in the Department of Economics at the University of Mannheim in Germany. Streb’s research focuses on the innovation history and regulation history of Germany in the nineteenth and twentieth century. He has contributed to the *Economic History Review*, *Explorations in Economic History*, the *Handbook of Cliometrics*, the *Journal of Economic History*, the *RAND Journal of Economics*, *Research Policy*, and other journals. In 2013, Streb published *Neue Deutsche Wirtschaftsgeschichte Des 20. Jahrhunderts* (German Economic History in the Twentieth Century), co-authored by Mark Spoerer.

Part I
Wars and State Formation in the Short and Long Run

Paths to Inclusive Political Institutions

Daron Acemoğlu and James A. Robinson

Abstract In this paper we present a new approach to thinking about the circumstances under which inclusive political institutions, consisting of a state with capacity and a broad distribution of political power, emerge. Different scholars have emphasized different paths towards such institutions, with some emphasizing modernization, and others emphasizing the necessity of state building as a prerequisite for democracy. We argue however, using the examples of Ancient Athens and Early Modern England, that inclusive political institutions emerge from a balanced increase in state capacity and the distribution of power. This path emerges in a particular basin of attraction. Though this basin depends on many parameters, we emphasize the crucial nature of informal institutions and social norms which put Athens and England onto this path. Outside of this basin a number of things may occur; social norms may be such as to stop a state forming, an outcome we illustrate with the Tiv of pre-colonial Nigeria; or when society is weaker a form of state formation can occur which creates a 'Paper Leviathan' which we illustrate with Colombia; finally when civil society is prostrate 'Despotic Leviathans' can be created, an outcome we illustrate with contemporary Rwanda. None of these latter paths lead to inclusive institutions or sustained prosperity.

Keywords Political institutions • State formation • Civil society • Social norms • Informal institutions

---

We thank Steve Hindle and Josh Ober for their invaluable comments and suggestions. This paper was written in honor of Mark Harrison at the time of his retirement from the University of Warwick. We have benefitted greatly from Mark's generosity and erudition, but most of all from his friendship and we look forward to a lot more of it.

D. Acemoglu
Department of Economics, Massachusetts Institute of Technology, E52-380,
50 Memorial Drive, Cambridge, MA 02142, USA
e-mail: daron@mit.edu

J.A. Robinson (✉)
Harris School of Public Policy, University of Chicago, 1155 East 60th Street,
Chicago, IL 60637, USA
e-mail: jamesrobinson@uchicago.edu

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*,
Studies in Economic History, DOI 10.1007/978-981-10-1605-9_1

JEL Classification: D7, H1, N4, P5

Introduction

What makes a society economically successful? Most social scientists would argue that the critical factors are the economic institutions, the rules that create patterns of incentives and opportunities in the economic sphere and which shape saving, investment and innovation. It is certainly true that economic institutions vary widely across societies both today and in history and are significantly correlated with economic performance (Acemoglu et al. 2014a, 2001; Acemoglu and Robinson 2012).

What creates variation in economic institutions? In Acemoglu and Robinson (2012), building on a great deal of earlier work stretching back at least to North and Thomas (1973), we argued that economic institutions have to be thought of as an outcome of political choices which are shaped by political institutions which influence how preferences are aggregated and the nature of incentives and constraints faced by those who exercise power. In this case, lying behind economic institutions that promote prosperity are political institutions that create particular economic institutions.

But what sorts of political institutions are associated with economic institutions that promote prosperity? In Acemoglu and Robinson (2012) we emphasized two dimensions of what we call "inclusive political institutions". First, there must be a state with capacity; second, political power must be broadly distributed in society.[1] We further emphasized the idea that transitions towards inclusive political institutions are the consequence of the mobilization of a "broad coalition" in society that, if it attains power, has the incentive to move on both margins, creating more state capacity and also making political power more broadly spread. This is certainly what happened after the Glorious Revolution in Britain in 1688 for example.

Many scholars however would argue that there were in fact basic incompatibilities between these two dimensions of inclusive political institutions. Huntington (1968) for example, claimed that if political participation expanded in the context of the lack of institutionalized modern political institutions, which centrally include

[1]There is a great deal of different terminology used in social science in this context. In Acemoglu and Robinson (2012) we used the terminology "political centralization" to refer to what we argued was the key aspect of having inclusive political institutions. Here that coincides with having state capacity by which we mean that the state develops some basic attributes, a monopoly of violence, a bureaucratic administration and fiscal system and has the "capacity" to provide public goods and regulate society and enforce laws. Some scholars would refer to such a state as "strong", though others would say such a state has "infrastructural power" and state strength is a different concept related to how autonomous the state is from society. In Acemoglu and Robinson (2012) we used the term pluralistic to refer to a situation where political power was broadly distributed since we wanted to emphasize that modern mass democracy was not necessarily either sufficient nor necessary for this.

the state, then the result was political instability and a society which was unable to provide high levels of well-being. He distinguished between "civic societies" which are "legitimate and law-abiding states, where rulers acted in the public interest" and "praetorian societies" which were "perverted or law neglecting systems, where the rulers acted in their own interests rather than those of the polity" (1968, p. 81). Once participation runs ahead of the institutionalization of political institutions and political instability starts it seems difficult to move towards a "civic society". He concludes "the experience of both early and late modernizers suggests that early attention to the problems of political organization and the creation of modern political institutions makes for an easier and less destabilizing process of modernization" (Huntington 1968, p. 399). Thus for Huntington, constructing a state with capacity must necessarily come before a broad distribution of power or the consequence is chaos and political instability.

Fukuyama (2011, 2014) likewise argues that to achieve something close of "inclusive political institutions" one must have the correct historical sequence with the rule of law being established first before state capacity and with democracy (related to our notion of a broad distribution of power) coming last. He views the rule of law as flowing from the great monotheistic religions, such as Christianity, and the emergence of state capacity as an elite project driven by inter-state warfare. Finally, democracy comes as a consequence of modernization and economic growth. Any deviation from this sequence leads away from inclusive institutions.

The approach of North et al. (2009) is to distinguish between 'natural states' which are organized to control elite sponsored violence, and 'open access societies' which are related to those we characterize as having inclusive institutions. The transition between these two regimes happens when some key "doorstep conditions" (if there is the rule of law for elites, if there are perpetually lived organizations and if there is civilian control of the military) are in place. Their approach again emphasizes that these things have to happen before real inclusion can take place. For example, the latter is dated to the middle of the 19th century in Britain, long after the doorstep conditions were secured.

In this essay we argue that while it is certainly true in some circumstances that it is difficult to create more state capacity and to make political power more broadly based in society at the same time, societies which have built inclusive political institutions have done precisely this. Yet they have not done it in the way suggested by the above scholars. In fact, we claim, once one looks closer at how states are built and how power is spread there is a basin of attraction in which these two processes are highly complementary. Characterizing the nature of this basin of attraction is the crucial task in understanding the emergence of inclusive political institutions. On the way to this goal one must revise much of the conventional wisdom on state formation in social science. The preponderance of this literature, which we survey later in section "The Academic Literature", emphasizes elite incentives to build or not build state institutions and finds successful societies emerge out of a constellation of parameters that induce elites to build state capacity, a process then followed by economic growth and ultimately broader political participation. According to this literature, states with capacity emerge then when it is in the interests of elites to create the necessary institutions, for example a modern fiscal system or a bureaucratized administration, usually when they are forced to do so to survive.

Yet this elite focused narrative, common to all the above scholars, gives a very poor account of the way state capacity and inclusive political institutions have actually formed historically. We illustrate this with case studies of the construction of the state in two very successful historical instances of inclusive institution building, Classical Athens, and Early Modern England. In both cases, the historical evidence suggests that popular pressure and involvement were absolutely critical in the emergence of inclusive political institutions. Elite interests were at play too, but they had to find an equilibrium with those of society.

To see the development of state capacity, or perhaps ‘state formation’, for example, simply as an elite project is like trying to analyze market equilibrium using only the supply function, without the demand function. The key insight here is that the emergence of inclusive political institutions is not a consequence of the emergence of elite created and controlled state capacity that then broadens political power (a rather unlikely scenario since the type of modernization mechanism typically appealed to by the above scholars is not supported by the empirical evidence, see Acemoglu et al. 2008, 2009). Nor is it generally the consequence of a society where political power is broadly spread which then builds the capacity of the state. Rather, it emerges from the coevolution of the state and society; in the successful cases we study both dimensions of inclusive political institutions emerge at the same time.

Though no doubt there are many parameters that influence the size and nature of the basin of attraction, in this essay we emphasize just one sub-set: the nature of informal institutions and social norms. Societies that lack states or have only nascent or incapacitated states are typically characterized by dense webs of informal institutions and social norms that make it very difficult to concentrate political power. When these norms are powerful and of a certain type, as for example in many pre-colonial societies in Africa or Melanesia, they stop any type of state emerging. Critically, such norms are not of a form which enables power, once accumulated, to be disciplined.

But this not always the case. As we show, in both the Athenian and English case, social norms facilitated the emergence of a capable state because they allowed people to be confident that once created the state could be controlled. In this case we argue a dynamic interaction between state and society can be triggered. In this dynamic the initial ability of society to control the state allows the state to emerge, increase its capacity and take on more tasks, but the formation of the state then crucially feeds back onto society reinforcing the conditions that gave rise to it. This allows the state to be further developed, again feeding back onto the nature of society. It is this dynamic that generates inclusive political institutions.

Finally, when societies lack informal institutions and norms that can block or discipline the accumulation of power, or they exist but are weak, they are unable to stop political entrepreneurs consolidating and centralizing power and building a state. Yet this path does not typically lead towards inclusive institutions because such a state can then use its capacity to oppose the spread of political power and the strengthening of civil society. In this part of the parameter space one finds the Leviathans which Thomas Hobbes believed were necessary to promote peace, security and prosperity.

Leviathans come in different types however; we emphasize the existence of 'Paper Leviathans' common to many post colonial states in Africa and Latin America, and 'Despotic Leviathans' such as the Communist state in China, or the Rwandan state. In both of these latter cases states formed without the effective cooperation or sanction of society and this has very important consequences for how they operate and behave. Nevertheless, such states possess some dimensions of capacity, for example a monopoly of violence and often administrative capacity, but because they operate without the input of society and lack accountability, they will not be able to provide the rule of law or inclusive economic institutions in a sustained way. The differences between Paper Leviathans and Despotic Leviathans are largely historical. Communist China, for instance, was able to utilize a 1,000 year history of bureaucracy and centralized institutions and identity. Though some places in Africa, like Rwanda, have access to some of these elements, such as a history of centralized state authority, which make for a Despotic Leviathans, most post-colonial states are at best Paper Leviathans.

These basins of attraction obviously have important consequences for economic development. On the one hand societies which have historically failed to develop states have remained poor because they have been unable to provide basic public goods. On the other hand, societies which have been able to build Despotic Leviathans, but in the context of what Scott (1998) calls a 'prostrate' civil society, have sometimes been able to generate what we called "extractive growth" (Acemoglu and Robinson 2012, Chapter 5). Yet, as we pointed out, such experiences of economic growth, such as in the Soviet Union, have been necessarily transitory. This essay makes a new argument in this respect. Since such states lack the cooperation of society and are built without any type of social consensus, they lack legitimacy and this further limits their potential for promoting economic growth even if this were their objective. It is rather in societies which build inclusive political institutions in which sustained economic growth emerges.

The paper proceeds as follows; in the next section we discuss the case studies of the rise of the Greek city state, particularly Athens, and the creation of the English state in the Early Modern period. Section "The Symbiotic Relationship between State and Society" then distills some lessons from these examples of the successful emergence of inclusive political institutions in particular emphasizing the way that state and society coevolved in the context of particular initial conditions with respect to informal institutions and social norms. In section "Outside the Basin of Attraction" we study societies which are outside of this basin of attraction in particular ways. We first examine the Tiv of pre-colonial Southeastern Nigeria which was a society which failed to create a state, at least partially because of the nature of their informal institutions. We then move to a different part of the parameter space where state formation can take place and capacity accumulated, but where similar mechanisms to those that operated in Tivland make the state weak in the sense that society does not wish to create an effective state. We show, using the example of Colombia, that such Paper Leviathans lack capacity for reasons which are 'from the bottom' (a la Tiv) and also 'from the top'. Section "Outside the Basin of Attraction" also examines a Despotic Leviathans, Rwanda, which governs without the cooperation of a prostrate civil society, but nevertheless because of historical factors is able to build and exercise capacity to the extent that it is capable of providing public goods and promoting economic growth (something difficult for Paper Leviathans). However, since Despotic Leviathans rule without social consensus, cooperation or accountability, they are just as likely to promote chaos and mass murder, something evident from the history of Rwanda (and also of course China). Section "The Academic Literature" then presents a survey of the literature on state building particularly focusing on the few studies which have identified some of the mechanisms which we believe are important for creating state capacity and trying to empathize what is distinct and new about our approach. Section "Conclusion" concludes.

The Basin of Attraction

The Emergence of the Athenian Polis

One of the most famous example of inclusive political institutions emerged in Greece from around 600 BC onwards, why? Bronze Age Greece certainly did not have inclusive institutions. The states of Knossos in Crete, or Mycenae on the mainland, were mostly ruled by warrior kings with little popular participation and based on command and control 'palace economies'. Yet these societies collapsed around 1200 BC, an event which ushered in the Greek Dark Ages. Population probably halved (Ober 2015) and social complexity greatly diminished.

From this collapse a very different type of society started to coalesce in Greece. Small relatively egalitarian chiefdoms emerged where elites and chiefs had little power and society developed informal institutions to discipline them. We can get a sense of what the political institutions of these societies looked like from Hesiod's *Works and Days* and Homer's description of Odysseus' Ithaca from the *Odyssey* (our discussion follows Morris and Powell 2010, Chapter 5).

Hesiod has a lot to say about the chiefs, called *basileis*, or more colorfully "gift-devouring *basileis*". In particular, he had no difficulty in taking them to task for being corrupt and not upholding the rule of law. Hesiod records that when his father died he and his brother Perses divided their inheritance in two but Perses bribed the *basileis* to get a larger share

> We already divided our inheritance, but you seized
> more than your share and held it, greatly praising
> the gift-devouring *basileis*, who like to take on a case
> like this. Fools! They know not that half
> can be more than the whole, and that great profit
> lies in a poor man's bread. (Hesiod, 37–41)

Odysseus had been the *basileus* in Ithaca before he disappeared for 20 years to take part in the Trojan Wars and then pursue a tortuous path back home. In the meantime most Ithacans thought he was dead and many aristocrats tried to woo his wife Penelope and grab his assets and position themselves to be the *basileus*. His son Telemachus protested and the way he did it is interesting. First, he did it at the assembly which could be summoned not just by the *basileus* but by citizens. During his appeal the ancient hero Aegyptius asks who has summoned the assembly “one of the youngsters? or one of the old-times?” so a broad section of Ithacan society could call the assembly. During his monologue Aegyptius clearly thinks of the “men of Ithaca” as a collectivity in charge of their society. Telemachus’ speech reveals that state institutions are weak and there is no police which can expel the intruders from Odysseus’ household. He appeals to the assembly to do this and implies their right to judge and discipline anyone in Ithaca, even those amongst those elites who are trying to take advantage of Odysseus’ absence. The egalitarian nature of the institutions is further illustrated when Zeus sends an omen in the form of two eagles flying in the sky above. But there is no monopoly in the interpretation of omens and no centralized control of religion, and the people in the assembly propose contradictory interpretations. Homer’s account portrays an Ithaca where there was a chief, but his power was checked by an assembly of citizens who were collectively able to discipline elites and even dictate what an acceptable social contract looked like. Though Homer is supposed to be describing Bronze Age society, it is more likely that his depiction of political institutions reflected those in Greece at the time he wrote, in the 8th century BC, in the Dark Ages.

During the Dark Ages a set of initial conditions seem to have emerged, similar to those Homer was describing in Ithaca, which allowed for the formation of the Greek polis and the subsequent coevolution of state and society. Critically, they reversed any concentration of political power that might have taken place in the hands of *basileis* or other elites. Morris (1987) showed that during this period grave goods declined and they became much more equally distributed. Ordinary people started to get proper burials and therefore the total number of burials increased dramatically. Before 700 BC a few people built very large houses, sometimes as much as 2,500 square feet but after this such mansions disappeared. Public displays of wealth seem to have diminished. For example, burials took place not within urban areas and under specific peoples’ houses, but in specific sites outside of towns where everyone, not just elites, were buried. Morris (1987, see also 1996) interprets this as the consolidation of a new egalitarian type of society which was very different from Bronze Age Greece.

Critically for the purposes of this paper, the construction of this new society allowed a particular type of far more inclusive polity to emerge. We can only observe some moments of this emergence, but these moments, best documented in Athens, show that the nature of the society and its broad distribution of power was critical for allowing a different sort of state to emerge with a great deal more capacity than before. This construction of this state then fed back onto the nature of society.

The most compelling evidence for this comes from the later institutionalization of informal institutions and the role they played in state formation. The two best examples of this come from the Athenian institutional revolutions of Solon in 594 BC and Cleisthenes in 508/07 BC. Solon's reforms mark the institutionalization of the Athenian state along an inclusive basis. No doubt there was conflict over this and the lack of consensus clearly showed afterwards with a transitory resurgence of tyranny, but just what Solon did is significant and left an enduring legacy. On the one hand he made enserfing an Athenian citizen illegal (there was a great deal of debt peonage at the time) and he implemented an egalitarian land reform (though just exactly what this entailed in not fully understood). So the institutional revolution featured movements towards far more inclusive economic institutions. This also included the abolition of what appear to have been restrictions on movement and location within Athens.

To lock into place the change in economic institutions Solon set a new political architecture for Athens. He divided the population into four classes based on their incomes from land. There was a popular assembly that all free (non-slave) Athenian men could attend, but state officials could only come from the highest three classes with only the richest being able to hold all political offices. Yet the poorest class, the *thêtes*, who were undoubtedly the majority of Athenians at the time, were nevertheless powerful. Though the 9 highest executive offices, the *archons*, had to be filled from the top two classes and the ex *archons* filled a council known as the Areopagus, their power was counterbalanced by juries which could hear appeals against their decisions. Any Athenian citizen could be a member of these juries and anyone could bring a suit infront of them. Moreover, the Assembly, of whom every male citizen was a member, elected the *archons* and made important decisions, such as going to war, democratically. The agenda for the assembly was drawn up by a Council of 400 equally representing the 4 traditional Athenian tribes, on which everyone was again represented. Solon himself observed, in a fragment which is preserved, that his institutional design was intended to create a balance of power between the rich and the poor

> …I gave to the *dêmos* as much reward as is fitting,
> neither taking away from their honor nor adding.
> As for those who had power and were admired for great wealth,
> I was careful that nothing improper happened
> to them. I took my stand, spreading out my strong
> shield over both parties, and not
> allowing either side to take unjust advantage. Solon fragments 4c, 5 (West)

One of Solon's most interesting laws was the Hubris Law (Ober 2005, Chapter 5). This forbade any act of hubris, behavior aimed at humiliation and intimidation, against any resident of Attica (the broader region in which Athens lay). Ober (2015, p. 150) notes "amongst the hubris' law's targets would have been rent-seeking elites who might have sought to mimic the Spartans by using threats and intimidation to reimpose limits on the free movement of poor citizens."

Solon's reforms did not stick, but they changed the way that Athenians thought about their political institutions and even the tyrants that followed had to appear to honor them. Ultimately they led to the reforms of Cleisthenes, which did stick. Cleisthenes was brought to power by a mass popular uprising against his opponents and their Spartan backers (according to one account amongst Solon's reforms was a law that required Athenians either to take sides when a civil conflict erupted or lose their citizenship when it was over, Ober 2015, p. 161). Cleisthenes started by re-organizing the basis of the state, abolishing the four tribes that had provided the people for Solon's Council of 400 and replacing it with a Council of 500 composed of people chosen by lot from ten new political units which were regionally based within Attica. To be on the council you had to be older than 30 but you could only serve for a year and at most twice in your lifetime which meant that most Athenian citizens served at some point in their life on this council. The president of this council was randomly chosen and served for 24 hours, but was in charge of the assembly of all citizens if it met. Hence the poorest person in Athens could be president of the Athenian assembly when it had to discuss a momentous decision.

From our perspective however one of the most telling things Cleisthenes did was to formalize the informal institution of 'ostracism'. The social norms that Homer suggested were part of the political equilibrium of Ithaca, involved the ability and legitimacy to sanction elites. This seems to have been the intent of Solon's Hubris Law. Cleisthenes legislated an institution which powerfully reinforced this equilibrium. Every year the Assembly could vote as to whether or not to ostracize someone. If at least 6,000 people voted and 50 % were in favor, then each citizen got to write the name of a person who they wanted ostracized on a shard of pottery. Whichever name got the most 'votes' was ostracized – banished from Athens for 10 years.[2] Ostracism was so potent that Themistocles, the genius behind the Athenians victory at Salamis over the Persians, and probably the most powerful man in Athens at the time, was ostracized for 10 years sometime around 476 BC when people began to worry that he was too powerful for the health of the state. Crucially, this was used very sparingly, only 15 people were ostracized over the 180 year period when the institution was in full force, but the threat of ostracism "off the equilibrium path" was a powerful way for citizens to discipline elites.

The emergence of inclusive political and economic institutions in Athens made the state far more powerful and unleashed a period of sustained economic growth (see Morris 2004; Ober 2015, Chapter 5). Athens created state institutions which could make effective collective decisions, provide public goods, such as the rule of law, and raise taxes. Indeed, it would have led to an Athenian super-state if this project had not been stopped by the Peloponnesian War.

The story about the emergence of inclusive political institutions and the Athenian state shows that this was not an example of elite driven statebuilding preceding democratization. Though reforms were legislated by Solon and Cleisthenes, they were institutionalizing and codifying a political equilibrium that already existed and at these moments the state developed capacity because people were confident that

[2]There are different interpretations of the origins and role of ostracism. We follow Morris (1987, 1996), Morris and Powell (2010) and Ober (2015) as seeing it as a tool via which citizens disciplined elites In her work Forsdyke (2000, 2005) has emphasized more the role of the institutions in resolving inter-elite contests.

they were able to control it, both democratically and through such institutions as the hubris law and ostracism. Solon's reforms in Athens reflected the fact that the elites already could not dominate society (witness the failure of Kylon to establish a tyranny Ober 2015, p. 148) and that there was a great deal of popular participation in government.

This perspective is driven home by the research of Gottesman (2014). Though we mentioned the Athenian state, in fact there was no professional bureaucracy or police force. Gottesman shows that to be implemented, laws passed by the Assembly had to generate a great deal of consensus more broadly in society and be implemented by popular force. 'Popular' included women, slaves and non-citizens. One reason this worked so well was the vibrancy of civil society. As Finley put it (1983, p. 82)

> This was not only a face-to-face society, it was also a Mediterranean society in which people congregated out of doors, on market days, on numerous festive occasions, and all the time in the harbour and the town square. Citizens were members of varied formal and informal groups – the family and household, the neighbourhood or village, military and naval units, occupational groups ... upper class dining-clubs, innumerable private cult associations. All provided opportunities for news and gossip, for discussion and debate.

Gottesman argues that there was a 'public sphere' in Athens which extended beyond even the hubris and ostracism laws where more subtle and routine ostracism and sanctions disciplined elites and that "one reason behind the strength of Athenian democracy was the fact that its leaders were constantly exposed to the ridicule of ordinary people" (Gottesman 2014, p. 19) and later he observes the presence of "insults and slights that would make life hard for anyone made it impossible for the politically active to pursue their ambitions" (p. 71).[3]

How did Athens, and more broadly the Greek city states get onto this path? Scholars have suggested several key reasons. One stemming from Childe (1942) and developed by Snodgrass (1980) is that the transition from the use of bronze to iron in itself redistributed political power in society. A Childe put it "cheap iron democratized agriculture and industry and warfare too". Both copper and tin were scarce and the use of bronze for weapons and armor encouraged, according to Childe (1942, p. 191), the concentration of political power. But iron ore, in contrast, was very common and the movement from bronze to iron took away the rents from elites who had previously dominated trade and made iron tools and weapons available to everyone at low cost. There were other important technological revolutions. One was the emergence of writing. Though Bronze Age Greece had forms of writing, known as Linear A and Linear B, they had been restricted to the elite and use primarily for record keeping by the state. Around 800 BC a new type of writing emerged which spread much more broadly in society (though of course literacy was low compared with modern societies). Other technological innovations included the perfection of hoplite warfare, perhaps connected to the spread of iron weaponry.

[3]See also Forsdyke (2008, 2012) on the bottom-up nature of the Athenian state and Ober (2012) for further analysis of social norms and democracy in Athens.

Polities who could amass more hoplites in battle had a military advantage and it is possible that this helped to underpin a further empowerment of the mass of citizens (see Morris 1987, Chapter 6 for scepticism about the importance of this). Finally, as is clear from Homer's discussion of Ithaca, political leaders could not claim to rule by divine right and there was no fusion between the political elite and religion. Religious power, such as that of the oracle at Delphi, was not controlled by political elites.

Thus there was technological, military and religious (ideological) change which pushed in the direction of a more egalitarian distribution of power. The research of Morris (1987) makes it clear that this path emerged in the context of Greek elites initially becoming less powerful and wealth becoming more broadly spread. But once this equilibrium was established the stage was set for the emergence of more elements of inclusive political institutions, particularly states and the accumulation of state capacity, because ordinary citizens had the tools to control them. States emerged and developed and power became more broadly distributed at the same time. Rather than state formation and a broad distribution of political power being inconsistent it was the latter that allowed the former to take place.

State and Society in Early Modern England

Sometime in December 1596 in Swallowfield, Wiltshire, a group of local people got together to compose a list of 26 resolutions which were to be the basis of local administration. These resolutions included monthly meetings (resolution number 25 – "the whole company promesethe to meete once in every monethe" – Hindle 1999, reproduces the entire text) with elaborate protocol (resolutions 1–3). For example, the first resolution read

> ffirst it is agre[e]d, That every man shal be h[e]ard at o[u]r metynge quyetly one after an other, And th[a]t non shall interrupte an other in his speeche, And th[a]t every man shal speake as he is fyrste in accompt, & so in order, th[a]t therby the depthe of every mans Judgment w[i]th reason may be concedered.

There was also to be bureaucratized record-keeping. Resolution 11 read

> And th[a]t ther be a paper Booke to Regester all o[u]r doynges & by or w[i]th [what] autorety or warrant wee do it consernynge her Ma[jes]ties service & one other Booke for the Churche & the poore.

The resolutions were mostly concerned with providing public order and counteracting "wilffull & vyle synns" (resolution 25) which ranged from fornication and illegitimacy (resolutions 8, 13); insubordination and disturbance of the peace (resolution 15); petty theft, malicious gossip, wood-stealing, pride, dissent, and arrogance (resolution 18); improvident marriage (resolution 20); harbouring inmates (resolution 21); profanation of the sabbath (resolutions 22, 24); and drunkenness (resolution 23).

Though they reveal the remarkable extent to which local communities in late Tudor England regarded themselves as self-governing, the resolutions and the meeting which wrote them did not drop out of the blue. They were an evolution from local manorial courts and eventually became institutionalized in the parish vestry, a meeting of local community members who met in the vestry of the church which was to form the backbone of local government in England until the 19th century.

Who wrote these resolutions? Not the local elite. Neither the two members of the local gentry Samuel Blackhouse and John Phipps who resided elsewhere, nor the local priest, who is referred to only once and then tangentially and in the third person. Hindle (1999) concludes that it was most likely that the resolutions were drawn up by a meeting of the 'middling sort' or person, the people who served as jurors, churchwardens, overseers of the poor, and local constables. These were not the rich of the community, since even leaving aside the two gentlemen, none of those who appear on jury lists between 1588 and 1613, and were therefore involved with the provision of these local public goods and plausibly a member of the group that wrote the resolutions, had enough income to be amongst the eleven taxpayers listed in the parliamentary lay subsidy tax return of 1594 (Hindle 1999, p. 843)

What was going on in Swallowfield in 1596 was almost certainly representative of late Tudor England and understanding it is crucial for understanding the emergence of inclusive political institutions in England.[4] On the face of it, this emergence has many similarities with that of Classical Greece, albeit with a 2,200 year time lag. The struggle between the Monarchy and Parliament which erupted into the Civil War of the 1640s led to a large movement towards more inclusive economic and political institutions. All domestic monopolies were abolished, the country became a republic and made significant strides towards building a modern state, for example with the introduction of the excise tax which was the main fiscal instrument for the next 200 years. But like Solon's reforms, those of Oliver Cromwell did not stick and the monarchy was restored in 1660 and launched again on the project of creating an absolutist state. It took the Glorious Revolution of 1688, like the reforms of Cleisthenes, to finally make the inclusive political institutions stick.

Yet as with the Athenian reforms, the institutional reforms of 17th century England built on a long history of the coevolution of state and society. The traditional story of the emergence of the modern English state dates it to the 'Tudor revolution of government' of the 1530s, first analyzed by Elton (1953) (see Coleman and Starkey 1986). Following the rise of the Tudor dynasty after the Battle of Bosworth in 1485 at the end of the Wars of the Roses, the aristocracy were gradually disarmed, a process which culminated in the 1558 Militia Act which incorporated the formerly liveried retainers of the aristocrats into the county militia under the control of a centrally appointed Lord Lieutenant of the County (Braddick 2000, Chapter 5). This is seen as progress towards a key plinth of the modern state, the

[4]Collinson (1994a,b) first drew attention to the significance of the Swallowfield resolutions, see the essays in McDiarmaid (2007) for discussion of his interpretation of them.

assertion of a 'monopoly of legitimate violence'. The reforms of the 1530s also began to separate the administration of the central state from the king's personal household, a step towards a modern bureaucracy and regional assemblies such as the Council of the North were abolished, leaving Parliament as the only representative institution. Finally, the 1530s also saw the break with Rome and the creation of the Church of England controlled by the Crown along with dissolution of the monasteries and expropriation of Church land (Heldring et al. 2015).

This gradual emergence and development of the modern English state has certainly been seen as a consequence of the decision of elites, Henry VII and Henry VIII and their powerful advisers such as Thomas Cromwell. It is true of course that these individuals did make critical decisions, but they did so in a very specific social context. For one, they needed the cooperation of society to implement any reforms or institutional changes, just as in Classical Athens. The autonomous people of Swallowfield had to buy into any project of state building and it had to take place in ways which respected their interests and demands. As Hindle (1999, p. 11) notes "Policies that rested on consensus were enforced, but only at a pace with which local governors were comfortable. Policies that they found dubious or that aroused violent opposition were quietly obstructed" (see also Fletcher 1986, p. 356).

At some level the reason for this is obvious. The lower levels of the English state were composed of exactly the same people who had composed the Swallowfield resolutions. Though they were unpaid, they were tasked with implementing state policy and providing public goods. Collinson (1994a, p. 25) proposes that participation was

> a capacity proved from day to day, year in and year out, by service on all kinds of juries, juries not merely to find a man guilty or innocent but to determine the responsibility for the clearing of a drain or the repair of a road or river bank; and in the time and effort spent in parish vestries, courts baron and courts leet, with all the powers to appoint officers, levy local rates, and fine and otherwise discipline their members.

Goldie (2001) estimates that in 1700 there might have been 50,000 parish officers at any one time in England, representing around 5 % of adult males, and since there was frequent rotation of offices the number of people who had held office must have been considerably larger. In 1800 he estimates the figure was more like 100,000 people. So the Tudor and Stuart states, like the Athenian one, were built upon the mass participation of the citizenry. In fact just as Ober (2015, p. 17) uses the phrase "collective self-governance by amateurs" to describe the participation of average Athenians in the state, so does (Hindle 1999, p. 24) too describe the English state as "imbued with the cult of the (often very experienced) amateur."

It wasn't just that the state needed the cooperation of society to implement its policy initiatives. Impulses and policies came from the bottom as well. This is best seen in the huge increase in litigation and demand for legal services that the state then provided. These were demanded by ordinary people who for instance were heavily involved in litigation even in supposedly royal dominated contexts such as the Star Court (Brooks 2009; Herrup 1989; MacFarlane 1981, 2013). It is also evident in the fact that many prominent pieces of legislation, such as the

Elizabethan Poor Law, was actually a local initiative (in Norwich) which was the picked up on by the central government Braddick (2000, Chapter 3). Indeed, the policy process in this period is summed up by Kumin and Wurgler (1997, p. 40) with the argument that it “appears more like a dynamic process of communication between center and localities rather than a one-sided drive towards ever greater penetration or acculturation.” Harriss (1993, p. 33) goes even further when he argues that government “was moulded more by pressures from within political society than by the efforts of kings or officials to direct it from above”.

An important question, in the light of our discussion of Athens, is why society was not threatened by this emerging strength of the state. One reason, as we saw, is that by embedding society into the state, the Tudor political elites could commit not to adopt policies which were inimical to the interests of the citizens. In addition, what we have seen is that policy initiatives were endogenous to the preferences and actions of people in society, not simply an outcome of elite projects. Hindle (1999, p. 16) characterizes this process by observing that “The early modern state did not become more active at the expense of society; rather it did so as a consequence of social need.” In addition, as Thompson (1971, 2001) and Wood (2013) have shown, rural England was characterized by a social equilibrium where ordinary people had strong views about appropriate behavior and resource allocation and were prepared to protest and riot in order to enforce these views. Moreover, they show that local elites respected these views and acted accordingly.

Of course some elite initiatives did threaten peoples’ interests or norms and in consequence they reacted violently. The Pilgrim of Grace of 1536 (Hoyle 2003) was a popular reaction to the creation of the Church of England, and Kett’s Rebellion of 1549 was a popular uprising as a consequence of the social changes which state formation was creating, particularly the rise of a new gentry class in the wake of the expropriation of monastic lands by Henry VIII (see Wood 2001, Fletcher and MacCulloch 2008). In the Northern Rising of 1569 the marcher Lords, the Percys and Dacres, rebelled against state formation, such as the Elizabethan Militia Act, which was stripping away their local authority. Yet as Wood (2001, Chapter 4) points out, by this time they had difficulty mobilizing traditional loyalties to mount an effective military challenge. Society has already changed and if the Percys and Dacres were opposed to these transformations, ordinary people were not necessarily. Moreover, this period saw a dramatic transition away from the type of open large scale rebellion which had characterized England for the previous centuries towards very different types of contestation (see Wood 2010, on this). One consequence of this can be seen by noting the difference between the War of the Roses of the 15th century and the English Civil War of the 1640s. The War of the Roses was a dynastic struggle between the House of York and the House of Lancaster over who was to control the English Crown. The Civil War was a conflict over the nature of political institutions and how society was to be organized.

Like Athens then, the emergence of the modern English state was based on the coevolution of state and society. England got onto a dynamic path towards a state with far more capacity than in the past and a broader distribution of political power, but as in Athens, the latter played a key role in the former.

We can also identify some of the factors which might have facilitated this particular path of institution building. One obvious one is the total collapse of labor coercion and feudal authority in England in the late Middle Ages. After the French Revolution of 1789 the National Assembly voted to abolish feudalism in France, but feudalism had been a dead letter in England for hundreds of years by that time. This seems to have allowed the emergence of a great deal of economic freedom and mobility in rural England. Also significant was the discovery of the Americas and the impact of Atlantic trade expansion on mercantile interests in a context where the Tudor and Stuart states were not powerful enough to create monopolies. This led to a very broad participation in trade and these new economic activities (Acemoglu et al. 2005a).

The Symbiotic Relationship between State and Society

The examples of the construction of inclusive political institutions in Classical Athens and in Early Modern England share some distinct features. In both cases the emergence of such institutions, in Athens with the reforms of Solon and Cleisthenes and in England of the Tudor period and then in the 17th century, combined both elements of inclusive political institutions – increasing capacity of the state and a broadening of political power. For example, the Glorious Revolution of 1688 built the capacity of state by leading to the bureaucratization of the fiscal system (Brewer 1990), but it also made society more democratic by institutionalizing regular parliaments and facilitating accountability through the use of petitions (Pincus and Robinson 2012). In neither case can the emergence of inclusive institutions be described as an elite project. Civil society played a critical role not just in demanding power but also in allowing the state to form, even demanding that it did form. We believe the major lesson of these two cases is that the state developed more capacity at the same time that civil society also became more organized and powerful.

The evidence suggests that a more subtle claim may be true. The capacity of the state and the organization of society fed on each other in a synergetic way. Tudor statebuilding, for example, was facilitated by the fact that civil society had social norms to discipline it. Society demanded that the state dispense justice and poor relief. Just as society impacted the state, so the state impacted society. As the state expanded, spreading law, infrastructure and poor relief down to the lowest level, people started to see themselves not simply as members of a local isolated community, but as part of a larger polity. This led them to reformulate their demands on the state. As state and society evolved in Tudor England both came to be re-defined. The change in the nature of conflict, for example, was a direct response to state formation. Though Thompson emphasized the stability of the rural social equilibrium, in fact the society he observed was undergoing significant changes. The growth of the state not only "drew together provincial communities into a more closely integrated national society" it "introduced a new depth and complexity to their local patterns of social stratification" (Wrightson 1982, pp. 222–223). State expansion changed society. With the decline of the sort of indirect rule via the Percys and Dacres which had characterized the political organization of the feudal English state and the consequent expansion of legislation, judicial activity and more intense social policy, people began to look not towards local elites, but towards the national government. Moreover, while Swallowfield might have regarded itself as politically autonomous, it had long been integrated into a national economy (Dyer 2007) a process which was now made more intense by state expansion (Hindle 1999, Chapter 2, see also Devereaux 2009). How this process changed society has been studied most carefully by Tilly (1995) for the 18th century.

By the 1570s and 1580s the older notion of the 'common weal' was replaced by the notion of 'public service' which Skinner (1978, Volume II, pp. 356–357) argues represented a shift towards a situation where the monarch should rule in the interests of all and that all subjects shared responsibility for the welfare of the state. Interesting, it is precisely in this period that the word 'state' takes on its modern connotation (see Skinner 1978, Volume I, pp. ix–x).

In the Athenian case there is a great deal of evidence that the reforms of Solon and Cleisthenes created a different sort of society, reinforcing the trends that gave rise to them. Though in both cases de jure political rights were restricted to citizens, which excluded women and slaves and non-citizens, as we mentioned above Gottesman (2014) has shown how the public sphere was much broader than this. Moreover, it got broader over time. He discusses the emergence of what he calls "mixed associations" which became institutionalized after 306 BC when a right of association emerged "for many groups that before could not express their solidarity publicly" Gottesman (2014, p. 50).[5] An earlier institutional innovation, which occurred between 353 and 330 BC was that of "supplication" whereby people had the right to petition the Assembly and ask for their action on a particular issue. This practice arose earlier but after this time fully 1/4 of Assembly meetings were given over to dealing with supplicants. Gottesman (2014, p. 103) surveying the existing inscriptions which resulted from these supplications concludes "they appear to involve only non-citizens".

These examples suggest that while it may be the case that in some circumstances the development of state capacity is difficult to reconcile with political power being spread broadly, in other circumstances they are highly complementary, indeed they feed on each other in a symbiotic relationship. A society which can wield power can allow the state to become more capable because it is confident it can control it. But as the state gets capacity and starts undertaking new functions it feeds back onto society, helping it to become more organized, coordinating its demands and thus becoming confident of allowing the state to develop even more capacity. This is a path which ultimately leads to inclusive political institutions.

[5] See Jones (1999) on the impact of Athenina state formation on the formation of associations.

Outside the Basin of Attraction

The Tiv

Studying ancient Athens or Early Modern Britain the main issue of interest seems to be what sort of state would emerge. But one should not take it for granted that the process which leads to a state would ever get off the ground. At the time of the Scramble for Africa, for example, about 1/3 of the people of West Africa lived in stateless societies (Curtin 1995, p. 399). Why did so many Africans live in societies without a state?

From an elite centric perspective this could only be because African elites did not find it worthwhile to build states given their understanding of the costs and benefits. Herbst (2000) for example, explains the relative absence of states in pre-colonial Africa as being a consequence of low population density implying that it was not worthwhile paying the fixed costs of state construction (see Osafo-Kwaako and Robinson 2013, for systematic evidence contradicting this hypothesis). From our perspective, however, an important constraint on state formation in large parts of the world historically was not the costs and benefits facing elites, but whether or not society was able to stop or control the process of state formation. This idea is well illustrated by the ethnographic literature on stateless societies all over the world. We focus here on one African example, the Tiv.

The Tiv are an ethnic group of Southeastern Nigeria who were stateless at the time of the colonization of the country, but nevertheless formed a coherent group with a well defined, even expanding (Sahlins 1961) territory. Historically they lived in villages of extended kin when the anthropologists Paul and Laura Bohannan studied them from the mid-1940s onwards (see e.g. Bohannan and Bohannan 1953). Bohannan (1958) recorded some of the social norms which kept the Tiv stateless. For example, during the summer of 1939 the colonial government and most social and economic activity came to a standstill in Tivland because of a cult called Nyambua. At the heart of the cult was a shrine and a man called Kokwa who sold charms to provide protection from mbatsav or "witches". Tsav means "power" in the Tiv language, particularly power over others. A person with tsav (it is a substance that grows on the heart of a person) can make others do what they want and kill them by using the power of fetishes. Crucially, although some people naturally have tsav, it can also be increased by cannibalism.

> A diet of human flesh makes the tsav, and of course the power, grow large. Therefore the most powerful men, no matter how much they are respected or liked, are never fully trusted. They are men of tsav – and who knows? (Bohannan 1958, p. 3)

The people with tsav belong to an organization – the mbatsav. Mbatsav has two meanings: Powerful people (it is the plural of tsav); A group of witches organized for nefarious purposes (robbing graves to eat the corpses). This is a pretty interesting double meaning. Imagine if in English the word "politicians" simultaneously meant "people who contest for or control elected government offices" and "A group of witches organized for nefarious purposes (robbing graves to eat the corpses)."

People initiated into the Nyambua cult were given a leather covered wand and a fly whisk. The whisk allowed one to smell out "counterfeit" tsav – created by cannibalism. In 1939 the whisks were pointed towards the 'chiefs' created by British indirect rule (the Tiv had no chiefs prior to the colonial period and in consequence the British colonial government imposed them from the outside). But Bohannan notes "the movement was not anti-British; it was anti- authority" (1958, p. 8). In fact, historical evidence and the oral history of the Tiv shows the cult had much deeper roots and was simply the most recent incarnation of a long series of anti-authority social movements

> When the land has become spoilt owing to so much senseless murder [by tsav] the Tiv have taken strong measures to overcome the mbatsav. These big movements have taken place over a period extending from the days of the ancestors into modern times (East 1939, p. 264).

In essence these religious cults were a way of stopping anybody becoming too powerful

> Men who had acquired too much power ... were whittled down by means of witchcraft accusations... Nyambua was one of a regular series of movements to which Tiv political action, with its distrust of power, gives rise so that the greater political institutions – the one based on the lineage system and a principle of egalitarianism – can be preserved (Bohannan 1958, p. 11)

But to have a state someone has to become powerful, start giving orders to others who accept their authority. Witchcraft accusations were therefore not just a method of stopping someone becoming too powerful but simultaneously stopped in its tracks a process that could have culminated in state formation. Hence the Tiv were a stateless society in the pre-colonial period.

Following Bohannan, our reading of the evidence is that the Tiv were kept stateless not because there was not a shortage of potential elites who wanted to accumulate power and start the process of state formation going. Indeed, Laura Bohannan, in her fictionalized recollections of fieldwork amongst the Tiv (published under the pseudonym, Eleanor Smith Bowen 1964) records several instances of Tiv chiefs attempting (unsuccessfully) to accumulate power. Rather, the Tiv had created a network of social norms and informal institutions which made it almost impossible for state formation to get started. The likely reason for this is that they feared that such an institution could not be controlled.

Critical for the present discussion is not simply the strength of these norms, as opposed to those of ostracism in ancient Athens, but the differing nature of these norms. In the Athenian case we showed how the nature of the norms allowed citizens to threaten elites "off the equilibrium path". Thus citizens could allow state formation to take place in the anticipation that they would be able to control it. Yet Tiv social norms were different. One reason for this was because

> Tiv ascribe all death to tsav. It is incorrect to say that tsav can cause death; rather, it wills the cause of death. "Power," in the form of tsav, is a source of volition. Death, like illness, does

> not have a single cause, or even a multiple cause in the way Westerners look at multiple causation. Rather, there is a cause and there is also a volition. Tiv tend to assign the same causes to death as we do-old age, accident, disease, and the rest. But knowing the cause is not sufficient for them. They must also know the source of the volition of the death. (Bohannan 1958, p. 4)

Since people are always dying, this means that the use of tsav and the desire to control it is every present in equilibrium. Another feature which stopped the Tiv using witchcraft threats "off the equilibrium path" was that the Mbatsav were thought to continually trick people into eating human flesh without realizing it, after which they could control an individual.

Cults like that of Nyambua were not the only way the Tiv blocked state formation, this also happened through their structure of age sets (see Smith Bowen 1964, and Bohannan's discussion of the Hoyo cult, 1958, p. 9). Our argument here is that the net effect of these was to push Tiv society outside the basin of attraction which allowed the processes we described in Classical Athens and Early Modern England to take off.

Leviathans

In historical Tivland a state building project could never get off the ground because people were afraid that the type of power concentration that it entailed could not be controlled. Such concerns are ever present in many parts of the world. However, in other cases civil society does not have the strength that it did in Tivland or instruments which are effective enough to stop a state forming. It may also be the case that in other contexts the benefits that a state can provide, in terms of public goods, outweigh the risks of creating it for society so citizens are willing to allow it. Maybe even more usual in modern history, states have been created by colonial powers in the formation of empires which they have then bequeathed to post-colonial societies. In Nigeria the British built a state apparatus of sorts which was then taken over by domestic actors after 1960. In this case the mechanisms which blocked the creation of a state in Tivland historically could not stop this happening, and nor were they well adapted to controlling such a colonial creation. Nevertheless, this did not mean that civil society recognized or accepted the legitimacy of such colonial constructions and this has been a great source of grievances and political instability in post-colonial Africa.

The fact that a project of state formation got off the ground in colonial and post-colonial Nigeria does not imply that this leads towards inclusive political institutions. Indeed, we now argue that when states form in situations where society is unable to control them then this leads to various types of 'Leviathans'. Such states can exercise authority and provide some types of public goods, but they lack the participation and cooperation of society and this limits what they can achieve. In some circumstances, while society cannot really control or stop a process of state formation, as in the above post-colonial societies, they may be able to withdraw from it and deny it legitimacy. This reticence of society is a common feature of what we shall call Paper Leviathans, which though they exist, are unable, and as we shall see unwilling, to exercise much power over society. We illustrate this with the case of Colombia.

A Paper Leviathan: Colombia

The Colombian state certainly exists, but it does dismally in undertaking all of the tasks that a state is supposed to undertake. It has never had the monopoly of violence in its territory and instead has conceded control of large swathes of territory to other armed groups. These include guerilla groups, such as the FARC (Revolutionary Armed Forces of Colombia) and the ELN (National Liberation Army) who have for 50 years engaged in massive extortion (to avoid being kidnapped or killed one pays "la vacuna" – the vaccination), kidnapping (see Grupo de Memoria Histórica 2013, on the extent of this), murder and massacres. They also encompass many types of paramilitary groups all the way to drug gangs and armed mafias.

The Colombian state doesn't just surrender the monopoly of violence, it surrenders state activities as well. To take one specific example (see Bautista et al. 2013), in 2006 President Uribe negotiated the demobilization of 34 Paramilitary groups. Around 30,000 people demobilized officially (probably a similar number just melted away). One of these groups was called the Peasant Self-Defense Forces of the Middle Magdalena. The roots of this group go back to 1977 when a peasant farmer, Ramón Isaza, formed a group of 10 men called the "Shotgunners" who took it upon themselves to fight back against the local expansion of the FARC guerilla group. Isaza quickly attracted the support of local landowners and elites and even drug dealers such as Gonzalo Rodríguez Gacha, one of the founders of the Medellín drug cartel. Yet Isaza was primarily interested in fighting communists and he ended up starting a war with the drug dealer Pablo Escobar. Though Isaza started off small, by 2000 he was running an army with 6 fronts controlling around 15,000 square kilometers. One of his key commanders was his son in law, Luis Eduardo Zuluaga (nicknamed "McGuyver" – McGuiver in Colombia – after the US Television character). McGuiver commanded 250 armed and uniformed men of the José Luis Zuluaga Front (FJLZ) whose territory included three core corregimientos (the main sub-municipality administrative unit) in the municipality of Sonsón (Jerusalén, La Danta, and San Miguel), but its power also extended to the rest of Sonsón, and into the neighboring municipalities of Argelia, El Carmen de Viboral, La Unión, San Francisco and San Luis and even as far as Communa 13 a suburb of Medellín. The FJLZ had a written legal system of 'estatutos' (statutes) that it tried to enforce and it had a rudimentary equality before the law in the sense that the same laws applied to members of the FJLZ as to the civilians. The FJLZ also had a bureaucratized organization with functional specialization between a military wing, civilian 'tax collectors' and a civilian 'social team' which appears to have been remarkably un-patrimonial. The front regulated trade and social life, had a mission statement, an ideology, a hymn, a prayer and a radio station called 'Integration in Stereo'. It gave out medals, including the "Order of Francisco de Paula Santander" and the "Grand Cross of Gold". The front taxed every landowner and businessmen in their territory. It even taxed drug-dealers and cocaine laboratories though it was not itself involved in the drug business. It also built extensive public goods including hundreds of kilometers of roads and the electrification of rural hamlets. It constructed schools and paid for teachers and musical instruments in others, it built a health clinic in La Danta, re-built an old-age people's home, built houses for poor people, started an artisan center, built sports stadia and a bull ring. All this in the full view of the Colombian state on the main road between the two biggest cities in the country, Bogotá and Medellín.

To get a sense of the extent to which Isaza and his commanders took on state like functions consider the following exchange between Isaza and the Magistrate in charge of his case under the Peace and Justice Law which President Uribe introduced to govern paramilitary demobilization

> Magistrate: Mister Isaza, do you recall if you received on any occasion some type of order so that the Self-Defense forces under your command, interfered in some type of election day or to enact some type of political intervention – for example, to influence people's decision to vote for some political party, movement or specific candidate?
>
> Ramón Isaza: Your Honor, we did not engage fully, for example, in activities such as gathering people for the elections; that was done mostly by the candidates directly. What we did do was in the veredas, such as La Danta, also in San Miguel or Cocorná which didn't have police, that were little towns removed from the main roads and there was no military or police force. There we protected these regions but we didn't tell anybody to vote for a particular person. Rather we looked after – what did we look after? – that maybe elections wouldn't be spoiled, that maybe fights or quarrels occurred. This we did in this and all the regions where these towns were; we provided security for the elections. (Fiscalia de Colombia 2012)

Thus the paramilitary forces took it upon themselves to make sure that elections were conducted properly.

The Colombian state doesn't just concede state like functions to paramilitary groups, it does so to guerillas as well (see León 2005, and Aguilera 2014, on the legal services provided by the FARC).

Just as the Colombian state has not established a monopoly of violence in its territory, neither has it developed a fiscal system to support a modern state. Tax revenues in Colombia are around 14 % of GDP today according to the World Bank, but this is still a remarkably small state and similar in relative size to that in Egypt, or Benin and Ghana in West Africa. Personal income taxes are just 1 % of GDP and the reliance on consumption taxes leads to a situation where the poorest decile of the population pay 4.5 % of their income in taxes while the richest decline pays 2.8 % (Joumard and Londoño Vélez 2013, Figure 4, p. 9). As one of the consequences, Colombia has the smallest number of government employees relative to the labor force of any Latin American country. The OECD (2013, p. 283) reports that government employees in Colombia were 4.7 % of the labor force compared to 15 % for the OECD average (OECD 2013, p. 283).

The state also fails in recruiting and promoting its employees meritocratically, a key part of state capacity. For example, according to the OECD (2013, p. 290) in some ministries 50 % of the employees are "provisional staff" who are recruited outside of the rules in place and are most likely patronage employments. As the OECD (2013, p. 291) puts it

> There is in effect a two-tier employment system in Colombia's civil service, with significant numbers of casual staff, hired on a discretionary basis by managers, working alongside tenured civil servants, often doing the same core work and often employed for considerable periods, but without security of employment or access to the terms and conditions of employment enjoyed by career civil servants.

The Colombian state lacks capacity in many other ways. It is incapable of conducting a regular national census, for example, something Britain has done since 1801 and the United States since 1790. Indeed the Colombian 'census' is not even a census in the strict sense since the government does not actually try to survey everyone except with respect to a few basic variables. The rest is a sample, not a census.

The absence of the monopoly of violence, lack of fiscal resources and bureaucratic capacity means that whatever well intentioned law is passed in Bogotá, it is very difficult to implement in much of the country. The state is structurally unable to provide basic public goods such as order or roads. It does not collect proper information on people or assets.

One of the most revealing illustrations of the Paper nature of the Colombian Leviathan comes from the implementation of President Santos' Victims Law. Signed into law in June 2011, this was his flagship policy aimed at restituting land to around 4.8 million internally displaced people who in the process of being displaced left behind 6 million hectares of land (about the size of Massachusetts and Maryland combined). The Victims' Law created an administrative and judicial process intended to return millions of hectares of stolen and abandoned land to displaced people over the course of a decade. At the time Human Rights Watch (2013) reported on it however, this policy had barely been implemented two years after the law passed and hundreds of people who have tried to use its procedures have been threatened (the law actually allows the state to pay for bodyguards for beneficiaries at risk, an admission that the Colombian state cannot guarantee security or provide basic public goods that one might have thought necessary to implement land restitution). One of the reasons why the bodyguards were needed was the extraordinary extent of impunity for those who displace people from their land. Human Rights Watch report that of the 17,000 accusations against the perpetrators of such violence, only 1 % have been prosecuted. As of June 2013 Human Rights Watch found the Restitution Unit had started to examine less than 20 percent of the more than 43,500 land claims it had received, and obtained rulings ordering restitution in roughly 450 of them. Just one family had returned to live on their land as a result of these rulings under the Victims Law. An update on the non-implementation of the Victim's Law was published in November 2014 by Amnesty International (Amnesty International 2014). By this date a little more than 300 people had had land restituted, though most did not actually ever get their land back since it is now occupied by "good faith" occupants, and 25 % of all the land restituted went to one person in the department of Meta! Thus for the 10 year duration which the law is supposed to have, at the rate it was being implemented at the time of the report, 1,300 Colombians will benefit in a country with 4.8 million displaced people. A revealing calculation made by the Fundación Forjando Futuro is that at the rate claims were being processes by the Unidad de Restitución de Tierras it would take 529 years to process all the existing claims. The Colombian state is therefore completely unable to implement even the policies which is prioritizes.

The Colombian state doesn't just neglect and ignore its citizens, it actively victimizes them. Evidence for this is the so-called 'false positives' scandal. When President Uribe was elected president in 2002 his mandate was to intensify counter-insurgency policy. In order to do this he introduced a series of high powered incentives for the military who could receive financial bonuses and holidays if they produced dead guerillas. We don't know exactly when these measures came into force, but leaked secret decrees from 2005 show they were certainly in force then. A consequence was that members of the army murdered and dressed up as guerillas possibly as many as 3,000 innocent civilians (a conservative estimate produced by an independent research institute is 1,500, see Acemoglu, Daron, Leopoldo Fergusson, James A. Robinson, Dario Romero and Juan F. Vargas, The perils of top-down statebuilding: Evidence from Colombia's false positives, unpublished, 2015, see also Human Rights Watch 2015). This experience led Colombian judicial prosecutors to refer to a military unit, the Batallón Pedro Nel Ospina, as a "group of assassins dedicated to creating victims to present them as having been killed in combat."[6] A shocking condemnation of the ill-discipline and lawlessness of the Colombian military and the incapacity of the state. More broadly the involvement of the army in the formation of paramilitary groups has now been well documented (e.g. Grupo de Memoria Histórica 2011; Ronderos 2014). Leaving aside the false positives, the Colombian army was indirectly responsible for hundreds of massacres, thousands of murders and hundreds of thousands of displacements.

The Colombian Leviathan is indubitably Paper. But why? We emphasize two classes of mechanisms, the first will be familiar from our history of the Tiv. There has been a general reluctance, which we call 'from the bottom', in Colombia since independence to cede power and authority to the central state the Colombians inherited from Spanish colonialism because of the fear that it cannot be controlled. But there has also been reluctance 'from the top'. By this we mean that those running this Paper Leviathan have been reluctant to launch a state building project not just because it runs into the opposition of society, but also because it turns out that there are advantages to elites from Paper Leviathans.

First reluctance from the bottom. The issues here are most clearly seen at the time of the 1863 Rionegro Constitution. This constitution introduced a sort of hyper-federal system which was an attempt to find an equilibrium between the constituent

[6]"un grupo sicarial dedicado a la consecución de víctimas para presentarlos como muertos en combate". http://lasillavacia.com/historia/el-batallon-que-gano-el-concurso-de-falsos-positivos-49218

states (we follow Fergusson, Leopoldo, Javier Mejia and James A. Robinson, Committing not to build a state: Theory and comparative evidence, unpublished, 2015, here). At the time there was a great deal of concern that one faction of the elite, possibly in the department of Cundinamarca, might launch a state building project or try to establish hegemony over the rest of the country. The response to this was to dismantle the national army and guarantee that the national state did not have the right to intervene in the affairs of the departments that constituted it. Indeed, Law 20 of 1867 declared that[7]

> Article 1. When in any State there arises a portion of citizens with the object of overthrowing the existing government and organizing another, the government of the Union shall observe the strictest neutrality between the belligerent groups.

The Colombian national state ("the Union" – "any State" refers to the individual federal states of Colombia) legislated its lack of the monopoly of violence. Though this law was repealed in 1880 in many ways the spirit of it lives on in Colombia and Fergusson, Leopoldo, Javier Mejia and James A. Robinson (Committing not to build a state: Theory and comparative evidence, unpublished, 2015) show that it was part of a commitment not to start a state building project. This commitment was cemented by a series of penal codes that made armed rebellion against the state an almost legal activity. These didn't just apply in the 19th century. As recently as 1980 the Colombian Penal Code contained the following clause:

> Title II Crimes Against the Constitutional Regime,
> Chapter I Of rebellion, sedition and riot
> Article 125. Rebellion. Those who by use of arms to overthrow the National Government, or who delete or modify the legal or constitutional regime by force, incur imprisonment of from three to six years.[8]

Hence armed rebellion against the government was punishable by three to six years in prison! The 1936 Penal code stipulated penalties of from sixth months to four years in prison. Though the 1980 version of this clause was repealed by Law 599 of 2000, 'political crimes' are still treated very leniently in Colombia. These laws and codes cemented a very decentralized political equilibrium in Colombia. In the agreement national elites could exercise little power over the constituent regional elites and thus had to negotiate to get anything done (what Robinson 2013, and González 2014, classify as a form of "indirect rule").

The persistence of this very decentralized nature of the Colombian state and how interests were opposed to the creation of a modern national state was starkly revealed by the legislative debates in the 1960s about the proposal to create a new agrarian reform institute. Conservative congressman Alvaro Gómez Hurtado opposed the initiative on the grounds that (24 January 1961):

---

[7]"Articulo 1. Cuando en algún Estado se levante una porción cualquiera de ciudadanos con el objecto de derrocar el gobierno existence y organizar otro, el gobierno de la Unión deberá observar la más estricta neutralidad entre los bandos beligerantes."

[8]ftp://ftp.camara.gov.co/camara/basedoc/codigo/codigo_penal_1980.html

> In my statement I analyzed some factors that make the Institute an eminently centralist organization, contrary to the country's reality which cries out for and rightly requires to be fundamentally decentralized, and I affirmed how the centralist trend may cause a series of tensions in the country, which can put the reality of agrarian reform on the road to a total or partial frustration. (Pinzón 1977, p. 258).[9]

The institute is not threatening simply because it is supposed to be in charge of land reform, it is also threatening because it represented a novel attempt to create a modern state, the creation of a partial 'Leviathan' as congressman Diego Tovar Concha put it

> And the creation of an institute of this magnitude naturally leads to the destruction of the ministries that may interfere with its activities. Therefore, senators, we are not being hysterical we are facing the real possibility of the creation of a leviathan. (Pinzón 1977, p. 289).[10]

This is something regional elites had fought against in the 1860s, and they were still fighting against it 100 years later.

So like the Tiv, Colombians have been reluctant to concede sufficient authority and power to the state. There is a state in Colombia, but it is Paper Leviathan.

The incapacity of the state in Colombia is not just a consequence of institutional moments such as the Rionegro Constitution. Informal institutions and social norms are important, just as they were with the Tiv and indeed Athens and England. In the Colombian case a nexus of social norms makes it difficult for the state to exercise authority, the most famous of which is actually an inheritance of the colonial system: "obedezco pero no cumplo" ("I obey but I do not comply") (see Melo 2012, on the importance of this in Colombian society). This adage was famously used by colonial elites in response to orders issued by Spain and it reflects a generalized antipathy to the implementation of state directives and laws. The disregard for the rules pervades all levels of Colombian society (Restrepo, Juan Diego. Medellín: ética y legalidad, unpublished, 2014, on Antioquia), but the important thing here is not to see them as a simple reflection of the lack of order, but rather as representing an antagonism to authority in the same way that the Nyambua cult represented such antagonism in Tivland in the 1930s.

Revealing evidence of the prevalence of such norms comes from an interview on La W radio station with a lawyer from the law firm of Brigard and Urrutia. This law firm is the go-to firm for the government in terms of legal advice, but in 2012 it turned out that they had been involved in the extensive creation of shell companies to game the land reform laws, allowing multi-nationals and some of the richest people

[9]"En mi exposición analicé algunos factores que hacen del Instituto una organización eminentemente centralista, contraria a la realidad del país que clama y exige con razón un descentralismo fundamental, y afirmé cómo esa tendencia centralista puede provocar en el país una serie de tensiones que por ser tensiones, pueden encaminar la realidad de una reforma agraria a una frustración total o parcial."

[10]"Y es que la creación de un instituto de esa magnitud, naturalmente lleva a su destrucción, de los Ministerios que puedan interferir con sus actividades. No es, pues Senadores, que nosotros estemos dando aquí un espectáculo de histeria frente a la posibilidad de la creación de ese leviatán."

in Colombia to illegally acquire vast tracts of land in the Eastern Planes to grow tropical palm and biofuels. The scandal caused one of the partners, Carlos Urrutia to resign as Colombian ambassador to the US. A journalist asked a lawyer from Brigard and Urrutia

> The question is: did you have to "stretch" the law so you could buy and keep the land?
> Brigard and Urrutia: The law is there to be interpreted. Here they are not white or black, they are there to be interpreted... we assumed one which we think is correct (interpretation of the law).[11]

In Colombia ('Here') the law is not black and white. One can interpret this testimony in different ways. Perhaps the most obvious is that individuals can often benefit personally by deviating from the law. But we would argue in this context that this is indicative of a deeper set of issues in Colombia. The antipathy to the rule of law reflects an antipathy to the state in Colombian society. Social norms help to limit state capacity and make sure that laws do not get enforced.

But this reluctance from the bottom does not exhaust the mechanisms which keeps the Colombian Leviathan Paper. Once such a state is set up it turns out that there are a great deal of mechanisms which inhibit those in control of the state from making it more effective (we follow, Acemoglu et al. 2013; Robinson 2013, 2016, see also Besley and Persson 2011). These include some obvious ones such as the fact that allowing Ramón Isaza to organize elections may lower the 'supply price of votes' in the sense that it is much easier and cheaper for political elites to buy votes from paramilitaries than it is to compete for votes themselves. There are a lot more subtle things going on, however. One mechanism relates to the symbiotic relationship between state and society we identified above. When a state lacks capacity and does not provide public goods, society is fragmented and if it can mobilize collectively it does so in a local way making parochial demands. Such demands are very easy for elites to deal with and defuse. Building a state would risk creating a very different sort of society which would be much more difficult to control. It is also clear in the Colombian case that the incapacity of the state in the sense of the rule of law and the absence of security allows elites to benefit by expropriating assets in the periphery, something the law firm Brigard and Urrutia were facilitating. So there is reluctance from the top as well as from the bottom.

The Colombian state then is a prime example of a Paper Leviathan. After independence from Spain, Colombian political elites created a state, or at least inherited one, but they were never able to agree on the creation of a modern state which had any real capacity. At the same time the state they created was such that even those who controlled the state and were in a position to at least attempt to build capacity had few incentives to change this situation. The result, from the economic point of view, is a situation of long-run economic divergence from countries which were able to construct inclusive political institutions.

[11] http://www.wradio.com.co/noticias/actualidad/abogado-de-la-firma-brigard--urrutia-rompe-su-silencio-en-la-w/20130614/nota/1915927.aspx

A Despotic Leviathan: Rwanda

We now turn to a different sort of Leviathan where civil society is not able to exercise the type of restraints we just saw in the Colombian case, let alone avoid the construction of a state, as in Tivland historically.

In 17th century a new state, the Nyiginya kingdom, emerged in what is now central Rwanda, created by a political entrepreneur called Ndori. He used two big tools to build a state, first, a political strategy: Clientelism or patrimonialism, and second, the development of a professional army which soon became hereditary. Ndori seems to have come from the north (he brought cultural items associated with southern Uganda) with a lot of cows, and he used the cows to build political alliances, lending them to people in exchange for their support – a system called ubuhake. As Vansina puts is

> Thus the cow as much as the bow and the spear founded the Nyiginya kingdom (2004, p. 47)

Ndori has a Hima, a name used widely in the region for people who looked after cows. These were not always high status people. In Buganda, to the north in modern Uganda, the Hima looked after the cows of the farming Buganda who regarded (and regard) them as menial people. Farmers had high status in Buganda. Within the Hima were the Tutsi, a sub-set which possibly related historically to a politically dominant sect, but with which the Nyiginya kingdom ended up identifying.

Existing chiefs had small armies of their kin and lineage members which they called up on an ad hoc basis. Ndori created institutionalized and named companies and armies under generals which had permanence and soon were to be allocated their own lands and herds of cows. The deepest effect of this new military organization was, according to Vansina, "the institutionalization of a glorification of militarism and martial violence that finally permeated the whole of Nyiginya culture as the armies became the foundation of the administrative structure of the realm" (Vansina 2004, p. 61). The Nyiginya professional army is very precocious. England did not have one until after 1688 and even then parliament had to vote to maintain it every year.

When Ndori constructed his state he was surrounded by other polities on an equal footing. In 1720 when Gisanura, one of his successors, was king of the Nyiginya kingdom, it was surrounded by the states of Nduga, Ndiza, Rukoma, Ruhanga and Marangara. As Gisanura continued the consolidation of the central state through the instrument of the army and ubuhake he also asserted ownership of all the land and all the cows in the kingdom. In the reign of Mazimpaka, Gisanura's successor, the army became hereditary, a unique event in East Africa, and the army commanders become the most important elite in the country. The army and the system of patrimonialism of ubuhake were the state. There was no bureaucracy. Even today Reyntjens (2015, p. 71) argues that "Rwanda is an army with a state rather than a state with an army."

Taxes were collected from farmers and herders in kind by the king and armies as they moved through the kingdom (the capital city only stopped moving and settled in Nyanza in the 1890s). The notion of a Hutu may have first emerged as a name for menials involved in supplying the army.

By the 1840s land seems to have started to become very scarce and a new type of reserved herding domain was created which were granted to well connected people. This took large amounts of land out of circulation. To allocate the increasingly scarce land two new types of local chiefs emerged called 'chiefs of the long grass' and 'chiefs of the land' and around 1870 a whole new intense system of exploitation was created by them – called uburetwa. Families had to deliver large proportions of the crops in kind as taxes and in addition spend 50 % of their time giving unpaid labor services.

In this period the elite became known as Tutsis. The meaning of Hutu spread so that it came to refer to farmers whether they were originally Hutu or not.[12] It was the chiefs of the long grass and of the land that institutionalized the distinction between Hutu and Tutsi in the context of the imposition of the uburetwa system which applied only to Hutus who had to be identified and singled out.[13]

Historically then a Leviathan with clear capacity in some dimensions, particularly with respect to coercion, emerged in Rwanda through the use of military power and patrimonialism which in the late 19th century was able to enserf most of the rural population and even impose an identity, Hutu, on them. In the 1950s the Tutsi monarchy was overthrown in the lead up to independence and the majority Hutu took over in the form of the Parmahutu ('Hutu Power') movement which eventually was taken over by Juvenal Habyarimana who established a one-party state under his dictatorship. Though the people in charge of the state changed, the capacity of the state remained. The micro-foundations of this have yet to be completely understood, but one hypothesis is that a long history of centralized (particularly militarized) authority socializes people into obeying state authority irrespective to who is in charge of the state (an argument made by Sebarenzi 2009). Reyntjens (2015, p. 25) makes the same observation in the post-Genocide period when he remarks

> An ancient state tradition played an undeniable role here: a mere two years after the extreme human and material destruction of 1994, the state was rebuilt. Rwanda was again administered from top to bottom.

The Hutu led state after independence promoted a very coercive model of economic development focused on forcing farmers to grow export crops like coffee (Verwimp 2013). Most significantly, state institutions played a central role in promoting the genocide of Tutsis in 1994. Des Forges (1999) documents in great detail the way that the genocide was planned in advance and orchestrated by state officials at all levels of the state (see Yamigizawa-Drott 2014, and Heldring 2014, for empirical evidence).

In 1994 the Hutu government collapsed and was replaced by the invading army of the Rwandan Patriotic Front (RPF) which has ruled the country ever since. The RPF and its military wing was mostly composed of Rwandan Tutsi refugees who had fled Hutu led violence and pogroms in the 1960s. The period since 1994,

---

[12]Vansina points out that there were many meanings to the term Hutu – all foreigners were called Hutus, for example.

[13]The creation of this system and the institutionalization of the Hutu/Tutsi distinction therefore clearly antedates the colonial period (though the Belgians almost certainly exacerbated it).

particularly since 2000 has seen a steady plan to consolidate the RPF and particular General Paul Kagame, their leader, into power in Rwanda. This plan has operated on many fronts. First, Rwandan democracy has been turned into a one-party state (Strauss and Waldorf 2011, and Reyntjens 2015, for overviews of the politics of post-genocide Rwanda). After an initial transitional government formed through negotiation, Kagame won the 2003 presidential election with 95 % of the vote and all parties represented in the legislature supported his nomination. This followed a process of local elections which used open voting and allowed the RPF to gain control over local elected authorities. Opposition parties were either banned or harassed and opponents murdered. In the 2010 election there was a 98 % turnout with 93.1 % of the vote for Kagame. In 2016 Kagame organized a referendum which voted to remove the presidential term limit.

Second, the state has been systematically packed with Tutsis. The United States Embassy (2008) reviewed 118 senior positions in ministries, parastatals and regulatory bodies in 2008, 2/3 were Tutsi and the memo concluded that "for all the government's exhortations to Rwandans to abandon ethnic identities . . . the political reality is self-evidently otherwise."

Third, civil society has been systematically crushed and repressed. International Crisis Group (2002) concluded that "the press, associations and opposition parties have been silenced, destroyed, or co-opted" (see also Amnesty International 2004, and Human Rights Watch 2004, on repression of civil society and Waldorf 2007, on the media).

Fourth, the authoritarianism of the regime spreads into the economy where the RPF now owns much of it (see Booth and Golooba-Mutebi 2011, and Gökgür 2012). Extractive political institutions tend to create extractive economic institutions.

The absense of constraining forces from civil society, of the sort we saw in Colombia, along with the history of militarism has allowed for the construction and maintenance of a Despotic Leviathan in Rwanda. This is compounded by the absence of the incentives 'from the top' which we saw have inhibited the development of capacity in Colombia. Without real elections there is no need to 'lower the supply price of votes' for instance. Much research details the capacity of the Rwandan state to coerce and monitor its people (see Purdekova 2011). An interesting example of this comes from the research of Sommers (2012, pp. 82–86) who shows how local officials kept files on every single person so that they knew "everything that people are supposed to be doing" something completely unimaginable in Colombia.

If it suits the regime, this state capacity can be used to some extent to provide public goods and promote development. But as Rwandan history so vividly shows, it can also be used to repress and terrorize its people. The interaction between coercion and economic development, so characteristic of previous Rwandan regimes,[14] is

[14]Reyntjens (2015) shows at many points the uncanny similarity between the pre-genocide politics and the post-genocide politics. For example, he notes how the Habyarimana years were always characterized by slogans of unanimity "all together for 100 %" when it came to elections, while

nowhere better seen than in the current government's rural development policy. This has involved the reorganization of the rural population and their forced re-location and simultaneously the forced consolidation of land and forced production of export crops after 2006 (see Ansoms 2009, and Reyntjens 2015, Chapter 7). But the fact remains that the capacity of the Rwandan state rests of its military dominance of society and is limited in many ways since it is not responsive to the preferences or demands of its citizens.

In trying to understand the Rwandan Despotic Leviathan it is interesting to mention that Vansina says that in Rwandan history he "observed the tendency of the rulers to resist any delegation of power both from excluding whole social groups from participation in the government and by eliminating their immediate competitors" (p. 202) The defining issue for him is: "How can one counteract the nefarious tendency whereby power is concentrated in the hands of a smaller and smaller group ... How to mitigate the alienation of the bulk of the population that such a concentration of power can engender?" Indeed.

Just as in the past the Rwandan Despotic Leviathan has the potential to impose huge costs on society. Of course, like all countries with extractive political institutions but some state capacity, it has the ability to provide some types of public goods (like order) and promote economic growth to some extent and it is this which has captured the imagination of a generation of Western politicians and aid workers so frustrated by the incapacity of African states. Yet our argument, and certainly the evidence from Rwandan history, is that contrary to what Hobbes argued, a Leviathan is just as likely to make life "nasty, brutish and short" as it is to remove such threats. In this vein Ingelaere (2010, p. 292) observing the similarity in the style of the development plans which characterized Habyarimana's state with that of Kagame notes "it was precisely a highly top-down, authoritarian, and non-democratic set of institutional structures and exercise of power that was of crucial importance in the administration of the genocide. Such forces are still present and potentially destructive."

Once the circumstances allow a Despotic Leviathan like this to be constructed the path towards inclusive political institutions is a difficult one.

The Academic Literature

Even though the notion of inclusive political institutions was introduced in Acemoglu and Robinson (2012) similar ideas have been extensively discussed in social science and history, though the two component parts are treated separately in two

today in Rwandan we have "Vote for Kagame 100 %" (p. 53). The continuities are even more disturbing than this. Reyntjens (2015, p. 31) quotes a 2003 speech of Kagame where he says "If they come with the objective of hindering our programs they will be injured ... Our clemency decreases ... To whoever prides himself of having harvested sorghum or maize, we will say that we have mills to crush them" – the sort of political discourse that brought Rwanda "cockroaches". See Desrosiers and Thomson (2011) for many connections.

very different literatures. One set of issues concerns the circumstances in which political power comes to be more broadly spread in society. This question has been studied most in the literature on democratization though the question of when a pluralistic society emerges, closer to our notion of a situation where power is broadly distributed, has been much less discussed. The other issue is that of state formation.

In terms of democratization there is currently a great deal of consensus that democratization comes as a result of pressure from below rather than something that is willingly created by elites. Though seminal work such as Moore (1966) proposed that democracy emerged from the presence of a strong middle class and other structural features such as the absence of labor repressive agriculture, recent research has been based on the arguments first developed by Therborn (1977) and Rueschemeyer et al. (1992) who provided case study evidence that democracy resulted more as a consequence of the demands of the, mostly poor, disenfranchised. This mechanism was first developed formally in Acemoglu and Robinson 2000, 2006, where we also present a great deal of case study and historical evidence. There is now also a body of econometric evidence supporting these ideas, e.g. Aidt and Franck 2015; Przeworski 2009 and Aidt and Jensen 2014, and it is consistent with the most robust facts about democratizations, such as the fact that they tend to follow economic crises (Brückner and Ciccone 2011).

Other arguments in the literature suggest that democracy emerges when elites give away power either because autocratic elites split and some decide that democracy is a better option than continued dictatorship (Collier 1999; O'Donnell and Schmitter 1986), or because democracy can be a way of resolving conflicts between different factions of elites, or because democracy solves a commitment problem that elites cannot otherwise solve (Lizzeri and Persico 2004). These arguments may certainly apply in some cases. For example, many Latin American countries adopted democratic political institutions and held many elections in the 19th century. One could certainly view these as ways for elites to allocate power (see Mazzuca and Robinson 2009, on the Colombian case). Yet these were typically riven with fraud and malpractices (Engerman and Sokoloff 2005) and far from representing the type of broad distribution of political power we have discussed in this paper. Modern democracy emerged only in the 20th century and again typically in the context of mass mobilization and demands for the excluded for political rights.

This is not to say that there are not other mechanisms that can help explain patterns of democratization. As Huntington (1991) first emphasized, democracy seems to come in waves and this is most likely caused by the diffusion of democracy (see Markoff 2014, and Acemoglu et al. 2014b, for econometric evidence).

The issue of where pluralism comes from has been much less studied. The seminal theoretical work on this is Dahl (1970) who argued that the pluralistic nature of US society was an important reason for its history of democracy. Why the US is pluralistic seems to be the idiosyncratic result of its history of colonization and frontier expansion (as emphasized by Turner 1921) which created the type of dense civil society studied in the early 19th century by de Tocqueville (2008). Putnam (1993) is perhaps the most important empirical study of pluralism which is closely connected to his characterization of Northern Italian society having high levels of social capital or a very dense 'associational life'. Putnam traces the roots of this to the medieval organization of Northern Italy with its free communes, city states and mercantile political dominance. Southern Italy, in contrast, suffered from a legacy of feudalism which created a non-pluralistic society with low levels of social capital. In Acemoglu and Robinson (2012) we argued that pluralism emerges from contestation with civil society playing an active role in demanding political change but only in the context where a 'broad coalition' makes these demands. In this essay we have advanced the hypothesis that the broad coalition is itself part of a dynamic co-evolution of state and society.

Our approach to the emergence of state capacity in this essay follows from our research on the emergence of democracy and pluralism in arguing that inclusive political institutions are rarely, if ever, willingly created by elites acting in the absence of pressure and demands from civil society. This applies as much to state capacity as to democracy or pluralism. Like our previous work this emphasis is heavily influenced by our reading of history and the research of historians as will be evident from the many citations particularly to the work of Ian Morris, Josh Ober, Michael Braddick and Steve Hindle. It is also heavily influenced by our reading of research in anthropology which has studied the great diversity of political institutions in human society and emphasized the importance of social norms and informal institutions in explaining these. The work of the Bohannans on the Tiv is seminal here and in Africa important work includes Evans-Pritchard (1940), the essays in Evans-Pritchard and Fortes (1940) and in Middleton and Tait (1958), see also Lee (1979). On Burma/Myanmar see Leach (1954) and much of the work of Melanesia has a similar emphasis, for example, Strathern (1975), Harrison (2006), and the synthesis of Flannery and Marcus (2014) who provide many other examples.

This emphasis is quite different from the preponderance of scholarship on the creation of state capacity which sits within a larger social science literature on 'state formation' which has been studied in political science, sociology, history and more recently economics. We cannot hope to do justice to it here, but instead point out the literature which is most closely related to our approach and what our contribution is relative to the main themes in the literature.

Scholarly research is focused on trying to explain the emergence of modern nation states in Europe since 1500 and their variants.[15] It is also dominated by scholars who see several large-scale structural factors as playing the key driving role. States are identified with some basic components, the monopoly of violence in a well defined territory, a centralized fiscal system and a bureaucratized

[15] A typical definition of a state in this literature would that of Mann (1984, p. 112) who argues that a state is:

1. *a differentiated* set of institutions and personnel embodying
2. *centrality*, in the sense that political relations radiate outwards from a center to cover a
3. *territorially demarcated area*, over which it exercises
4. a monopoly of *authoritative binding rule-making*, backed up by a monopoly of the means of physical violence.

administration, and state formation is measured by movement in any of these three dimensions. These coincide with the notion of state capacity we have used. Of this the fiscal side has received the most attention. For example, Brewer (1990) provided a seminal account of the rise of the excise tax system in Britain, and empirical work has focused on tracking the rise and centralization of tax revenues (Dincecco 2009, 2011; Karaman and Pamuk 2013; O'Brien 2011).

The literature has emphasized several structural factors as driving European state formation (possibly in interaction with each other); inter-state warfare; trade and mercantile expansion; the collapse of medieval society and feudalism; and the rise of capitalism.

The hypothesis that inter-state warfare drove the emergence of the modern state and the accumulation of state capacity was proposed originally by Hintze (1975) and was elaborated by Bean (1973) and Tilly (1975, 1990). According to this hypothesis the military revolution of the 17th century (Roberts 1956) forced states to build modern fiscal systems because warfare became much more expensive. Without a fiscal system to fund armies a country could not survive inter-state competition. As a consequence kings and elites were forced to develop fiscal systems to survive. This theory is widely accepted by many scholars (e.g. Ertman 1997; Fukuyama 2014; Herbst 2000; Mann 1986, 1993).

Nevertheless, there were large differences between the different types of state that emerged in Europe in the Early Modern period. Some, like Britain or the Netherlands, were constitutional, others, like France or Prussia were absolutist. Some, Britain and Prussia, were much more bureaucratized (with high capacity), others, like France, were much more patrimonial (lower capacity) in the way the state was organized. To account for these facts, theories emphasize the interaction between warfare and other factors. Hintze (1975) emphasized geography, claiming that because Britain was an island and not subject to invading armies, this allowed the state to become constitutional (rather than absolutist). Tilly (1990) distinguishes between places which were able to tax mercantile wealth and take the 'capital intensive path' as opposed to others which had to adopt a 'coercion intensive path'. In the former, Britain, constitutional rule emerged while in the latter, Prussia, bureaucratic absolutism emerged. Mann (1986, 1993) focuses on two sorts of state power; 'despotic', by which he means the extent to which the state could formulate objectives and policy without the input of society; and 'infrastructural' power which is the ability to penetrate society and implement policy (close to our notion of capacity as we mentioned above). The British state in the Early Modern period, for example, had low despotic power but high infrastructural power. Mann notes that "societies are constituted of multiple overlapping and intersecting sociospatial networks of power" (1986, p. 1). State formation involves asserting autonomy from such networks and exerting control over them. For Mann this process is driven by the fact that states provides some collective advantages in terms of public good provision,[16] and that they more effectively use force and thus better enable communities to survive (e.g. Mann 1984, pp. 119–120). For example, he attributes the rise of states in Early Modern Europe to inter-state warfare and mercantile and capitalist expansion which generated a demand for public goods only the state could provide. What type of power the state then ends up possessing, despotic or infrastructural, then depends on a host of idiosyncratic factors. For example, in Britain where there was a strong mercantile class, kings had to negotiate with them to raise resources and taxes and this reduced the potential for accumulating despotic power. Finally, Ertman (1997) argues that while inter-state warfare is important, its timing is crucial. European states that experienced warfare early on, circa 1450, were more likely to develop patrimonial, not modern bureaucracies. Moreover, he also stresses the interaction between warfare and different histories of local government (see also Downing 1992, on this). For example, in the British case the combination of constitutional rule and bureaucratization came from the juxtaposition of late warfare and the history of strong autonomous local government rooted in the Anglo-Saxon state. This latter was important because it led to the particular form of regional representation in parliament (as opposed to one based on 'estates') and created a legislature which felt it had greater rights over the determination of policy.

The 'war made states' literature therefore has already combined some of the other 'forcing variables' which have been claimed to create states. Mercantile expansion following the discovery of the Americas features heavily in Tilly's theory, and economic growth and the demand for public goods features centrally in Mann's theory. Other scholars, particularly Spruyt (1994, 2009) place economic expansion more directly at the heart of their theory emphasizing that state formation was a process of bargaining between political and economic elites with the different outcomes being determined by heterogeneous economic opportunities of different political organizations (contrast the unity of England with the very fragmented political context in France).

Another class of theories of European state formation see it as a class project to create more capacity to repress and discipline society in the context of the collapse of feudalism as a result of trade expansion and the Black Death (Anderson 1974; Hechter and Brustein 1980, and see Acemoglu, Daron, James A. Robinson and Ragnar Torvik. The Political Agenda Effect and State Centralization, unpublished, 2015). This argument is related to those of Braddick (2000) about the role of the English civil war in inducing several projects to develop state institutions, particularly fiscal ones, and Slater (2010) and Saylor (2014) have both argued that state formation can be induced by the desire to develop tools and resources to repress domestic opponents and challengers.

A final class of theories relates state formation to ideological change and the Reformation and or the Enlightenment. Gorski (2003), for instance, sees the

[16] As he puts it (1984, p. 135) "autonomous state power is the product of the usefulness of enhanced territorial centralization to social life in general."

emergence of modern states coming as a response to the need to discipline society following the Reformation, while Mokyr (2009) argues that the enlightenment is the pivotal moment which leads to the economic and institutional transformation of modern Europe. Such views are shared by many historians, e.g. Israel (2013).

A recent literature in economics has tried to investigate formally many of these ideas. Acemoglu (2005) constructed a model in which a self-interested ruler taxes and invests in public goods and citizens make investment decisions. Lack of state capacity is detrimental to economic development because it discourages the ruler from investing in public goods as he anticipates that he will not be able to raise taxes in the future. Besley and Persson (2009, 2011) also emphasize the importance of state capacity and suggest that developing it will be deterred when groups that hold power are afraid that the state they build will be used against them in the future. Acemoglu et al. (2011) and Acemoglu et al. (2013) provide various models of persistence of states with low state capacity and Gennaioli and Voth (2015) develop a model of the interaction between warfare and state capacity (see Thies 2007, 2005, for some econometric evidence). Other formal analyses are Mayshar et al. (2011) and Acemoglu et al. (2015).

For our purposes here the main drawbacks of all such approaches are clear from the discussion of the last few sections. First of all, to take the 'war made states' thesis, it is an elite-centric theory of state formation where fiscal systems, for example, emerge as a result of an elite cost benefit analysis. Even though Ertman (1997) does allow the autonomy of parliament, rooted ultimately in civil society, to play a role, he still portrays the process of state formation as something driven by kings, though moulded in a particular way by the power of the legislature.[17] This hypothesis does a very bad job of explaining the Greek or Early Modern English evidence we discussed above (see Pincus, Steven C. A. and James A. Robinson, A Theory of state formation in early modern Europe, unpublished, 2015, on the lack of explanatory power of the 'war makes states' hypothesis in the English case and Centeno 2003; Kurtz 2012; Soifer 2015, for its problematic application in Latin America). We saw there that state formation was the result of the inter-play between elites and society and driven both by the fact that both sides saw advantages in the provision of various types of public goods and that society felt confident in its ability to control a state which had greater capacity. This created the impulse to build the institutions to provide them. This is a very different path to the constitutional and bureaucratic state in Britain than the one envisioned by Tilly or Mann.

Indeed the approach of this literature has been heavily shaped by the notion of state autonomy (see the introduction of Evans, Rueschemeyer and Skocpol 1985, for an assertion of this view) where the state and those running it take on a life of their own outside of the control of social actors. Many scholars argue that state autonomy is almost a necessary condition for successful economic development

[17]Ertman also focuses heavily on the development of the Medieval English state and sees the Early Modern period on which we focus as characterized by a general disintegration of state capacity. The opposite of our analysis.

(see the discussion in Barkey and Parikh 1991). Our view is radically different to this. We argue that it is actually impossible for an autonomous state to have capacity other than in very narrow dimensions, like the Rwandan case we discussed, because the input of society is critical for building state capacity. Our theory does allow for an autonomous state to generate growth, but only transitorily.

Scholars have also tried to theorize about the interaction of state and society. This surfaces a little in theories such as Spruyt's based on negotiation between political and economic elites, but this is again an elite discussion and the negotiation does not shape the nature and interests of the participants. Mann also sees society as continually in contestation with the state, sometimes benefitting from it and sometimes trying to capture it. Mann similarly recognizes that differences in social networks have important consequences for how the state forms, for example they explain why the French state took a much more patrimonial form in the Early Modern Period (engaging in venality and selling offices) while the British state did not because it was better able to extricate itself from society.

Reacting to what he regarded as too much emphasis on state autonomy, Evans (1995) coined the phrase 'embedded autonomy' to refer to a state which was autonomous from society, but at the same time sufficiently embedded within it that it understood the problems and needs of society. In his theory however, society plays little role in shaping the nature of the state, and the state does little to shape society. Related work is due to Migdal (1988, 2001) who emphasized the difficulty for a state to become autonomous from a 'strong society' with 'weak states' being those which were captured or dominated by society. Arguments of this type appear in the literature on the history of the state in the US where the state is argued to be weak because society was organized and had access to democratic institutions before the state was created (Skowronek 1982).

Though the example of the Tiv is related to the idea that a strong society may stop state capacity emerging, we emphasize instead the complementarities between the development of society and the state and how they feed on each other and how this is critical in understanding the emergence of capable states (part of inclusive political institutions).

The literature on state formation has of course extensively discussed the impact of the state on moulding and transforming society and creating new identities with Weber (1976) being the canonical reference (see also Gellner 2009). Most relevant to our discussion, this perspective has been extended to social movements. In the literature states may influence social movements by offering resources that they can try to control (Tilly 1978), or by favoring one group rather than another, for example through selective policy or repression (McAdam et al. 1988).

Our emphasis on how the state shapes society is related to the work of Habermas (1989) who saw the origins of the 'public sphere' to be an inclusive place in society where people come together to discuss and deliberate and form opinions. Though Habermas viewed this as in a sense the outcome of state formation, noting that "Civil society came into existence as the corollary of a depersonalized state authority" (1989, p. 19) yet in fact his main argument is that it is the rise of the bourgeoisie and economic and social change which created the public sphere in Early Modern Europe.

> In its clash with the arcane and bureaucratic practices of the absolutist state, the emergent bourgeoisie gradually replaced a public sphere in which the ruler's power was merely represented before the people with a sphere in which state authority was publicly monitored through informed and critical discourse by the people, (Habermas 1989, xi).[18]

de Tocqueville (1856) also argued that state formation changed the nature of society and that (pp. 101–102) a consequence was

> to powerfully assimilate the French people. National unity loomed through the surviving distinctions of rank. The laws were uniform. As the eighteenth century advanced, the numbers of edicts, declarations, and Orders in Council, which applied the same rules with equal force to all parts of the kingdom, became larger and larger. Subjects as well as rulers entertained ideas of a general uniform system of legislation that should bear equally on all ... not only have all the provinces grown like each other, but the men also. A marked resemblance began to exist between men of all ranks and stations.

Nevertheless, he saw the project of state building as entirely elite driven and discusses no feedback from this new society to state formation. Relatedly other scholars, particularly Skocpol (2003) have emphasized how the nature and extent of social capital in society is critically related to the behavior and policies of the state (see also Rothstein and Stolle 2008).[19]

None of this research has placed the same emphasis on the mutual dependence of civil society on the state and how paths towards inclusive political institutions are only made possible by having a civil society which has instruments to discipline the state.

As we noted in the introduction, our arguments are related to the "sequencing debate" in political science: does the state come before democracy, as Huntington argued, or democracy before the state. Not all political scientists have accepted Huntington's position. Carothers (2007) provides a series of arguments against it, in particular that once a non-democratic state is constructed there is no necessity that it will foster democracy, noting that "state-building beyond the initial stage is best pursued at the same time as democratization, with an effort to find points of complementarity and mutual reinforcement" (p. 20). Similarly Mazzuca and Munck (2014) suggest many ways in which democracy is consistent with state building and argue that the preponderance of evidence is against the "state first, democracy second" thesis. Our arguments are complementary to both of these papers, but we emphasize fewer and different mechanisms and try to place this debate into a larger conceptual framework.

Another difference between the social science literature on state formation and our essay is that the literature focuses on societies where in a sense there was already a well defined state. It may have been true in France in 1700 that the authority of

---

[18]See the essays in Lake and Pincus (2007) and Condren (2009) for the evidence on the nature of the public sphere in Early Modern England.

[19]Other authors who have identified the impact of the state on the nature of society include Katznelson (1985) who argued that it was the organization of the US state that determined why working class social movements took the form they did and see Birnbaum (1981, 1988) for other relevant examples.

Louis XIV was undermined by regional assemblies and by the fiscal and institutional fragmentation of the country (Collins 2009). But there was nevertheless a well defined state and scholars have focused on how this made itself more capable (e.g. through the appointment of intendants in the provinces to implement policy). In this essay, with our discussion of the Tiv and its relationship to Greece and England, we have implicitly tried to make a link to the literature on 'pristine state formation' by anthropologists and archaeologists. This literature, which we have been heavily influenced by, studies the forces which lead a stateless society to construct a state. This set of issues is typically regarded as something distinct from the standard social science literature on state formation a position we disagree with. Interestingly, though this literature also emphasize structural features such as population density, trade potential and warfare (see Johnson and Earle 2000) it has also emphasized that social norms and informal institutions, particularly egalitarian ones, represent a big impediment to the early stages of state formation, particularly the creation of chiefdoms (see the examples in Clastres 1989, and Flannery and Marcus 2014, and our citation above). It has also presented a plethora of arguments about why such norms break down or are overcome (see Flannery and Marcus 1996, for an example based on the theory of Friedman 1998). To our knowledge however this literature has not argued systematically that variation in informal institutions are a key to understanding different patterns of the dynamics of state capacity. Our emphasis on social norms and informal institutions as impediments to state formation and as influencing how states work in Africa is however consistent with a recent literature by archaeologists on the history of the state in Africa (Dueppen 2014; McIntosh 1988; Monroe 2012; Monroe and Ogundiran 2012).

Our research has also been heavily influenced by the work of James Scott. Our analysis of the implications of a Despotic Leviathan is related to Scott's (1998) arguments about when states create socially disastrous projects. He has also argued (e.g. Scott 2010) that people often see the formation of states as fundamentally inimical to their interests and that they therefore resist the process of state formation. We agree with this position but argue that people only resist the process of state formation to the extent that they anticipate not being able to control or influence the state. We have provided examples of people demanding that the state provide services and public goods in this context, something not possible in Scott's theory.

Conclusion

In this essay we have presented a new approach to thinking about the emergence of inclusive political institutions. In Acemoglu and Robinson (2012) this is seen to be a consequence of the coalescence of a broad coalition who are the losers from a system of extractive institutions. If this coalition can solve the collective action problem then it can take power and create inclusive institutions; both a state that has capacity and a broad distribution of political power in society.

But as the work of other scholars suggests, the component parts of inclusive political institutions may not necessarily be compatible with each other. A state with capacity may block expanded political participation and a society with an extensive distribution of political power may not find it easy to build a capable state. Moreover, just why is it that some societies develop broad coalitions pushing for institutional change and others do not?

In this essay we have argued that under some circumstances there is a basin of attraction where the two dimensions of inclusive political institutions are highly complementary. Indeed, they feed on each other to create a particular dynamic which leads to inclusive political institutions with a pattern of mutually reinforcing feedback. We argue that understanding this basin of attraction is a key to understanding the emergence of inclusive political institutions. Though in general the parameter space is multi-dimensional in this essay we have focused on one type of distinction which we believe is critical for determining the potential of different societies to move inside the basin of attraction: the strength and nature of social norms and informal institutions.

Though our emphasis is consistent with the current scholarly consensus on democratization, its argument runs counter to the great preponderance of scholarship on state formation which has taken an elite centric position and has seen the creation of modern states through the lens of a cost benefit analysis of elites. The demands or nature of civil society rarely feature in these calculations. In addition, scholars such as Huntington (1968) and Fukuyama (2011, 2014) emphasize a particular path towards their vision of inclusive political institutions which first involves state building and only later democracy, or institutions where political power is broadly spread. Though we present no econometric test in this paper we used two case studies from the history of classical Athens and Early Modern England to show that this sequenced elite centric approach cannot explain their transitions towards either state capacity or inclusive political institutions more broadly.[20]

The sequencing view we critique is similar to the one argued against by de Tocqueville in *The Old Regime and the French Revolution* when he claimed that the French reformers of the 18th century, such as the Physiocrats, were mistaken when they "sought reforms before liberties" and intellectuals such as Quesnay were wrong when they argued that "The system of counterpoises is a fatal feature of government". Rather political liberties are a critical complement to reforms and building the state, and one cannot rely on automatic processes such as modernization to subsequently bring "liberty". The Physiocrats, like many modern scholars,

[20]Though we do not have the space to go into this here, in fact the historical evidence from much of Western Europe supports a similar interpretation there, see Lenman and Parker (1980), the essays in Blickel (1989) and Blockmans et al. (2013), and Wheeler (2011) and Sreenivasan (2013) on Germany. The Swiss case is perhaps the most obvious one where inclusive political institutions were constructed from the bottom up. Another rather obvious case is the United States (recall de Tocqueville 2008) at least after one moves beyond simplistic ideas about the role of great men like James Madison detached from their societies.

proposed that one should rely on education to make sure that state promoted social interests and Quesnay even claimed that "Despotism is impossible in an enlightened nation". The aim of such reformers therefore was "not to destroy, but to convert the absolute monarchy" (de Tocqueville 1856, p. 194). Our econometric analysis elsewhere (Acemoglu et al. 2005b) supports the position of de Tocqueville when he says "Such was the literary nonsense they wanted to substitute in the place of political guarantees" (p. 194). Like his, our reading of the historical evidence is that the "state first, democracy later" development path is not a development path at all.

The ideas proposed in this paper help to clarify just where the broad coalition of Acemoglu and Robinson (2012) comes from. In the classical Athenian and Early Modern English case, the answer presented here is that it came from a pattern of social norms and informal institutions which facilitated not just the formation of the state, but also led to a distinct strengthening of civil society. In a sense, the broad coalition which overthrew the government of James II and the Stuart regime in 1688 was the consequence of the way that the English state had formed in the previous 150 years. Ironically even the Stuart state had helped to create the society which overthrew it and then transformed it in a more inclusive direction.

Such an argument may be though unsatisfying in the sense that it pushes the explanation for differences in political and economic development paths further back in time. Nevertheless, as we argued in Acemoglu and Robinson (2012), divergent development historically is the result of institutional differences which start out small but cumulate over time. England did not experience the industrial revolution because of some huge shock to the society in the eighteenth century. It did so because of a long process of institutional change in which both the state and the society coevolved and entered into a virtuous circle ultimately leading to the broad coalition that overthrew James II in 1688. England did not get onto this path because it was a radically different society from other Western European societies in the Middle Ages. But small differences mattered and it was inside a basin of attraction which turned out to have profound consequences.

References

Acemoglu, Daron. 2005. Politics and economics in weak and strong states. *Journal of Monetary Economics* 52: 1199–1226.

Acemoglu, Daron, Francisco A. Gallego, and James A. Robinson. 2014a. Institutions, human capital and development. *Annual Reviews of Economics* 6: 875–912.

Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson. 2015. State capacity and economic development: A network approach. *American Economic Review* 105(8): 2364–2409.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91: 1369–1401.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2005a. The rise of Europe: Atlantic trade, institutional change, and economic growth. *American Economic Review* 95(3): 546–579.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2005b. From education to democracy? *American Economic Review* 95(2): 44–49.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. Income and democracy. *American Economic Review* 98(3): 808–842.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2009. Reevaluating the modernization hypothesis. *Journal of Monetary Economics* 56: 1043–1058.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2014b. Democracy does cause growth. Forthcoming in the *Journal of Political Economy*.

Acemoglu, Daron, and James A. Robinson. 2000. Why did the west extend the Franchise? Growth, inequality and democracy in historical perspective. *Quarterly Journal of Economics* CXV: 1167–1199.

Acemoglu, Daron, and James A. Robinson. 2006. *Economic origins of dictatorship and democracy*. New York: Cambridge University Press.

Acemoglu, Daron, and James A. Robinson. 2012. *Why nations fail*. New York: Crown.

Acemoglu, Daron, James A. Robinson, and Rafael J. Santos-Villagran. 2013. The monopoly of violence: Evidence from Colombia. *Journal of the European Economic Association* 11(s1): 5–44.

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni. 2011. Emergence and persistence of inefficient states. *Journal of European Economic Association* 9(2): 177–208.

Aguilera Peña, Mario. 2014. *Contrapoder y justicia guerrillera : fragmentación política y orden insurgente en Colombia, 1952–2003*. Bogotá: Instituto de Estudios Políticos y Relaciones Internacionales, Universidad Nacional de Colombia: Debate.

Aidt, Toke S., and Raphael Franck. 2015. Democratization under the threat of revolution: Evidence from the great reform act of 1832. *Econometrica* 83(2): 505–547.

Aidt, Toke S., and Peter S. Jensen. 2014. Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820–1938. *European Economic Review* 72: 52–75.

Amnesty International. 2004. Rwanda Deeper into the Abyss – Waging War on Civil Society.

Amnesty International. 2014. A land title is not enough. https://www.amnesty.org/en/documents/AMR23/031/2014/en/

Anderson, Perry. 1974. *Lineages of the absolutist state*. London: Verso.

Ansoms, An. 2009. Re-engineering rural society: The visions and ambitions of the Rwandan elite. *African Affairs* 108(431): 289–309.

Barkey, Karen, and Sunita Parikh. 1991. Comparative perspectives on the state. *Annual Review of Sociology* 17: 523–549.

Bautista, María Angélica, Juan Sebastián Galan, Juan Diego Restrepo, and James A. Robinson. 2013. Acting Like a State: The case of the Frente José Luis Zuluaga. Research in progress.

Bean, Richard. 1973. War and the birth of the nation state. *Journal of Economic History* 33(1): 203–221.

Besley, Timothy, and Torsten Persson. 2009. The origins of state capacity: Property rights, taxation and politics. *American Economic Review* 99(4): 1218–1244.

Besley, Timothy, and Torsten Persson. 2011. *The pillars of prosperity*. Princeton: Princeton University Press.

Birnbaum, Pierre. 1981. State, centre and bureaucracy. *Government and Opposition* 16(1): 58–77.

Birnbaum, Pierre. 1988. *States and collective action: The European experience*. New York: Cambridge University Press.

Blickel, Peter, ed. 1989. *Resistance, representation and community*. Oxford: Clarendon Press.

Blockmans, Wim, and André Holenstein, and Jon Mathieu, eds. 2013. *Empowering Interactions: Political Cultures and the Emergence of the State in Europe 1300–1900*. Burlington: Ashgate.

Bohannan, Paul. 1958. Extra-processual events in Tiv political institutions. *American Anthropologist* 60: 1–12.

Bohannan, Paul, and Laura Bohannan. 1953. *The Tiv of Central Nigeria*. London: International African Institute.

Booth, David, and Frederick Golooba-Mutebi. 2011. Developmental patrimonialism: The case of Rwanda. African power and politics working Paper No. 1

Bowen, Eleanor Smith. 1964. *Return to laughter: An anthropological novel*. New York: Anchor.

Braddick, Michael. 2000. *State formation in early Modern England, c. 1550–1700*. New York: Cambridge University Press.

Brewer, John. 1990. *The sinews of power: War, money and the English State, 1688–1783*. Cambridge: Harvard University Press.

Brooks, Christopher W. 2009. *Law, politics and society in early Modern England*. New York: Cambridge University Press.

Brückner, Markus, and Antonio Ciccone. 2011. Rain and the democratic window of opportunity. *Econometrica* 79(3): 923–947.

Carothers, Tom. 2007. The sequencing fallacy. *Journal of Democracy* 18(1): 12–27.

Centeno, Miguel. 2003. *Blood and debt: War and the nation-state in Latin America*. College Station: Penn State University Press.

Childe, V. Gordon. 1942. *What happened in history*. London: Penguin Books.

Clastres, Pierre. 1989. *Society against the state*. Cambridge: MIT Press.

Coleman, Christopher, and David Starkey, eds. 1986. *Revolution reassessed: Revisions in the history of tudor government and administration*. New York: Oxford University Press.

Collier, Ruth B. 1999. *Paths towards democracy: The working class and elites in Western Europe and South America*. New York; Cambridge University Press.

Collins, James B. 2009. *The state in early Modern France*, 2nd ed. Cambridge: Cambridge University Press.

Collinson, Patrick. 1994a. *De republica anglorum*: or history with the politics put back in, in his *Elizabethan essays*. London: Bloomsbury Academic.

Collinson, Patrick. 1994b. The monarchical republic of Queen Elizabeth I. in his *Elizabethan essays*. London: Bloomsbury Academic.

Condren, Conal. 2009. Public, Private and the idea of the ‘Public Sphere’ in early–Modern England. *Intellectual History Review* 19(1): 15–28.

Curtin, Philip. 1995. The European conquest. In *African history: From earliest times to independence*, ed. Philip Curtin, Steven Feierman, Leonard Thompson and Jan Vansina. New York: Pearson.

Dahl, Robert A. 1970. *Polyarchy*. New Haven: Yale University Press.

Des Forges, Alison. 1999. *Leave none to tell the story: Genocide in Rwanda*. New York: Human Rights Watch.

Desrosiers, Marie-Eve, and Susan Thomson. 2011. Rhetorical legacies of leadership: Projections of ‘Benevolent Leadership’ in Pre- and Post-Genocide Rwanda. *Journal of Modern African Studies* 49(3): 429–453.

Dueppen, Stephen A. 2014. *Egalitarian revolution in the Savanna: The origins of a West African political system*. New York: Routledge.

Devereaux, Simon. 2009. The historiography of the English state during ‘the Long Eighteenth Century’: Part I – Decentralized perspectives. *Historical Compass* 7(3): 742–764.

Dincecco, Mark. 2009. Fiscal centralization, limited government, and public revenues in Europe, 1650–1913. *Journal of Economic History* 69: 48–103.

Dincecco, Mark. 2011. *Political transformations and public finances*. Cambridge: Cambridge University Press.

Downing, Brain M. 1992. *The military revolution and political change*. Princeton: Princeton University Press.

Dyer, Christopher, ed. 2007. *The self-contained village?: The social history of rural communities 1250–1900*. Hereford: University of Hertfordshire Press.

East, Rupert. 1939. *Akiga’s story: The Tiv tribe as seen by one of its members*. London: Oxford University Press.

Elton, Geoffrey R. 1953. *The tudor revolution in government: Administrative changes in the reign of Henry VIII*. New York: Cambridge University Press.

Embassy of the United States. 2008. Ethnicity in Rwanda – who Governs the Country? US Embassy in Kigali, Ref. Kigali 480, August 5th, Wikileaks http://wikileaks.org/cable/2008/08/08/KIGALI525.html

Engerman, Stanley L., and Kenneth L. Sokoloff. 2005. The evolution of suffrage institutions in the New World. *Journal of Economic History* 65(4): 891–921.

Ertman, Thomas. 1997. *Birth of the leviathan: Building states and regimes in medieval and early Modern Europe*. New York: Cambridge University Press.

Evans, Peter B., Dietrich Rueschemeyer, and Theda Skocpol, eds. 1985. *Bringing the State back in*. Chicago: University of Chicago Press.

Evans, Peter B. 1995. *Embedded autonomy*. Princeton: Princeton University Press.

Evans-Pritchard, E.E. 1940. *The nuer*. Oxford: Clarendon Press.

Evans-Pritchard, E.E., and Meyer Fortes, eds. 1940. *African political systems*. New York: Oxford University Press.

Finley, Moses I. 1983. *Politics in the ancient world*. New York: Cambridge University Press.

Fiscalia de Colombia. 2012. Audiencias de control de legalidad de los cargos. February 16th 2012. Morning session.

Flannery, Kent, and Joyce Marcus. 1996. *Zapotec civilization: How urban society evolved*. London: Thames and Hudson.

Flannery, Kent, and Joyce Marcus. 2014. *The creation of inequality: How our prehistoric ancestors set the stage for monarchy, slavery, and empire*. Cambridge: Harvard University Press.

Fletcher, Anthony J. 1986. *Reform in the provinces: The government of Stuart England*. New Haven: Yale University Press.

Fletcher, Anthony J., and Diarmaid MacCulloch. 2008. *Tudor rebellions*, 5th rev. ed. New York: Routledge.

Forsdyke, Sara. 2000. Exile, ostracism and the Athenian democracy. *Classical Antiquity* 19: 232–263.

Forsdyke, Sara. 2005. *Exile, ostracism and democracy: The politics of expulsion in ancient Greece*. Princeton: Princeton University Press.

Forsdyke, Sara. 2008. Street Theater and popular justice in ancient Greece. *Past and Present* 201: 3–50.

Forsdyke, Sara. 2012. *Slaves tell tales: And other episodes in the politics of popular culture in ancient Greece*. Princeton: Princeton University Press.

Friedman, Jonathan. 1998. *System, structure, and contradiction: The evolution of 'Asiatic' social formations*, 2nd ed. New York: AltaMira Press

Fukuyama, Francis. 2011. *The origins of political order: From prehuman times to the French revolution*. New York: Farrar, Straus and Giroux.

Fukuyama, Francis. 2014. *Political order and political decay: From the industrial revolution to the globalization of democracy*. New York: Farrar, Straus and Giroux.

Gellner, Ernest. 2009. *Nations and nationalism*, 2nd ed. Ithaca: Cornell University Press.

Gennaioli, Nicola, and Hans-Joachim Voth. 2015. State capacity and military conflict. *Review of Economic Studies* 82(4): 1409–1448.

Gökgür, Nilgün (2012) "Rwanda's ruling party-owned business enterprises: Do they enhance or impede development? IOB Discussion Paper 2013.03, Antwerp.

Goldie, Mark. 2001. The unacknowledged republic: Officeholding in early Modern England. In *The Politics of the Excluded, c 1500–1850*, ed. Tim Harris. Basinstoke: Palgrave.

González, Fernán E. 2014. *Poder y Violencia en Colombia*. Bogotá: CINEP.

Gottesman, Alex. 2014. *Politics and the street in democratic Athens*. New York: Cambridge University Press.

Gorski, Philip S. 2003. *The disciplinary revolution: Calvinism and the rise of the state in early Modern Europe*. Chicago: University of Chicago Press.

Grupo de Memoria Histórica. 2011. *Silenciar la Democracia. Las Masacres de Remedios y Segovia*. Bogotá: Grupo de Memoria Histórica.

Grupo de Memoria Histórica. 2013. *Una Sociedad Secuestrada*. Bogotá: Grupo de Memoria Histórica.

Habermas, Jürgen. 1989. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge: MIT Press.

Harrison, Simon J. 2006. *Stealing people's names: History and politics in a Sepik river cosmology*. New York: Cambridge University Press.

Harriss, Gerald. 1993. Political society and the growth of government in late Medieval England. *Past and Present* 138(1): 28–57.

Hechter, Michael, and William Brustein. 1980. Regional modes of production and patterns of state formation in Western Europe. *American Journal of Sociology* 85(5): 1061–1094.

Heldring, Leander. 2014. State capacity and violence in Rwanda. https://sites.google.com/site/leanderheldring/

Heldring, Leander, James A. Robinson, and Sebastian Vollmer. 2015. Monks, gents and industrialists: The long-run impact of the dissolution of the English Monasteries. NBER Working Paper #21450.

Herbst, Jeffrey I. 2000. *States and power in Africa*. Princeton: Princeton University Press.

Herrup, Cynthia B. 1989. *The common peace: Participation and the criminal law in seventeenth-century England*. New York: Cambridge University Press.

Hindle, Steve. 1999. Hierarchy and community in the elizabethan parish: The swallowfield articles of 1596. *The Historical Journal* 42(3): 835–851.

Hindle, Steve. 2002. *The state and social change in early Modern England, 1550–1640*. New York: Macmillan.

Hintze, Otto. 1975. In *Historical essays of Otto Hintze*, ed. F. Gilbert. New York: Oxford University Press.

Human Rights Watch. 2004. Rwanda: Parliament seeks to abolish rights groups. https://www.hrw.org/news/2004/07/02/rwanda-parliament-seeks-abolish-rights-group

Human Rights Watch. 2013. The risk of returning home. http://www.hrw.org/sites/default/files/reports/colombia0913webwcover.pdf

Human Rights Watch. 2015. On their watch: Evidence of senior army officers' responsibility for false positive killings in Colombia. http://www.hrw.org/report/2015/06/24/their-watch/evidence-senior-army-officers-responsibility-false-positive-killings

Hoyle, Richard W. 2003. *The pilgrimage of grace and the politics of the 1530s*. New York: Oxford University Press.

Huntington, Samuel. 1968. *Political order in changing societies*. New Haven: Yale University Press.

Huntington, Samuel. 1991. *The third wave*. Norman: University of Oklahoma Press,

Ingelaere, Bert. 2010. Peasants, power and ethnicity: A bottom-up perspective on Rwanda's political transition. *African Affairs* 109(435): 273–292.

International Crisis Group. 2002. Rwanda at the end of the transition: A necessary political liberalization.

Israel, Jonathan I. 2013. *Democratic enlightenment: Philosophy, revolution, and human rights, 1750–1790*. New York: Oxford University Press.

Johnson, Allen, and Timothy K. Earle. 2000. *The evolution of human societies: From foraging group to agrarian state*, 2nd ed. Stanford: Stanford University Press.

Jones, N. F. 1999. *The associations of classical Athens: The response to democracy*. New York: Oxford University Press.

Joumard, Isabelle, and Juliana Londoño Vélez. 2013. Income Inequality and Poverty in Colombia – Part 2. The Redistributive Impact of Taxes and Transfers. http://www.oecd-ilibrary.org/economics/income-inequality-and-poverty-in-colombia-part-2-the-redistributive-impact-of-taxes-and-transfers_5k487n4r0t8s-en

Katznelson, Ira. 1985. Working class formation and the state: Nineteenth Century England in American Perspective. In *Bringing the state back in*, ed. Peter B. Evans, Dietrich Rueschemeter, and Theda Skocpol. New York: Cambridge University Press.

Karaman, K. Kıvanç, and Şevket Pamuk. 2013. Different paths to the modern state in Europe: The interaction between domestic political economy and interstate competition. *American Political Science Review* 107: 603–626.

Kumin, Beat, and Andreas Wurgler. 1997. Gravamina, Petitions and early modern legislation in England and Hessen-Kassel. *Parliaments, Estates and Representations* 17: Kindle Edition.

Kurtz, Marcus J. 2012. *Latin American state building in comparative perspective*. New York: Cambridge University Press.

Lake, Peter, and Steven Pincus, eds. 2007. *The politics of the public sphere in early Modern England*. Manchester: Manchester University Press.

Leach, Edmund E. 1954. *Political systems of highland Burma: a study of Kachin social structure*. Cambridge: Harvard University Press.

Lee, Richard B. 1979. *The !Kung san: Men, women and work in a foraging society*. New York: Cambridge University Press.

Lenman, Bruce, and Geoffrey Parker. 1980. The state, the community and the criminal law in early Modern Europe. In *Crime and the law: A social history of crime in Western Europe since 1500*, ed. V.A.C. Gatrell, Bruce Lenman and Geoffrey Parker. London: Macmillan.

León, Juanita. 2005. *País de Plomo*. Bogotá: Editorial Aguilar.

Lizzeri, Alessandro, and Nicola Persico. 2004. Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's "Age of Reform". *Quarterly Journal of Economics* 119(2): 707–765.

Mann, Michael. 1984. The autonomous power of the state: Its origins, mechanisms and results. *Archives Européennes de Sociologie* 25: 185–213.

Mann, Michael. 1986. *The sources of social power: Volume 1, A history of power from the beginning to AD 1760*. New York: Cambridge University Press.

Mann, Michael. 1993. *The sources of social power: Volume 2, The rise of classes and nation-states, 1760–1914*. New York: Cambridge University Press.

Markoff, John. 2014. *Waves of democracy: Social movements and political change*, 2nd ed. New York: Routledge.

Mazzuca, Sebastián L., and Gerardo Munck. 2014. State or democracy first? Alternative perspectives on the state-democracy nexus. *Democratization* 21(7): 1221–1243.

Mazzuca, Sebastián L., and James A. Robinson. 2009. Political conflict and power-sharing in the origins of Modern Colombia. *Hispanic American Historical Review* 89: 285–321.

McAdam, Douglas, John D. McCarthy, and M.N. Zald. 1988. Social movements. In *The Handbook of Sociology*, ed. Neil J. Smelser. Beverly Hills: Sage.

McDiarmaid, John F., ed. 2007. *The monarchical republic of early Modern England: Essays in response to Patrick Collinson*. Aldershot: Ashgate.

MacFarlane, Anthony. 1981. *The justice and the Mare's Ale: Law and disorder in seventeenth-century England*. New York: Cambridge University Press.

MacFarlane, Anthony. 2013. *Law and violence in an Elizabethan village*. CreateSpace Independent Publishing Platform.

Mayshar, Joram, Omer Moav, and Zvika Neeman. 2011. Transparency, appropriability and the early state. CEPR Discussion paper 8548.

McIntosh, Susan K., ed. 1988. *Beyond chiefdoms*. New York: Cambridge University Press.

Melo, Jorge Orlando. 2012. Debería haber una ley. http://www.jorgeorlandomelo.com/ti_deberia.html

Middleton, John, and David Tait, eds. 1958. *Tribes without rulers*. New York: The Humanities Press.

Migdal, Joel. 1988. *Strong societies and weak states: State-society relations and state capabilities in the third world*. Princeton: Princeton University Press.

Migdal, Joel. 2001. *State-in-society: Studying how states and societies transform and constitute one another*. New York: Cambridge University Press.

Mokyr, Joel. 2009. *The enlightened economy*. New Haven: Yale University Press.

Monroe, J. Cameron. 2012. Power and agency in precolonial African states. *Annual Review of Anthropology* 42: 17–35.

Monroe, J. Cameron, and Akinwumi Ogundiran, eds. 2012. *Power and landscape in Atlantic West Africa*. New York: Cambridge University Press.

Moore, Barrington. 1966. *The social origins of dictatorship and democracy*. Boston: Beacon Press.

Morris, Ian. 1987. *Burial and ancient society: The rise of the Greek City State*. New York: Cambridge University Press.

Morris, Ian. 1996. The strong principle of equality and the archaic origins of Greek democracy. In *Demokratia: A conversation on democracies, ancient and modern*, ed. Joshua Ober, and Charles Hedrick. Princeton: Princeton University Press.

Morris, Ian. 2004. Economic growth in ancient Greece. *Journal of Institutional and Theoretical Economics* 160(4): 709–742.

Morris, Ian, and Barry B. Powell. 2010. *The Greeks: history, culture and society*, 2nd ed. New York: Prentice Hall.

North, Douglass C., and Robert P. Thomas. 1973. *The rise of the western world*. New York: Cambridge University Press.

North, Douglass C., John Wallis, and Barry R. Weingast. 2009. *Violence and social orders: A conceptual framework for interpreting recorded human history*. New York: Cambridge University Press.

Ober, Joshua. 2005. *Athenian legacies: Essays in the politics of going on together*. Princeton: Princeton University Press.

Ober, Joshua. 2012. Democracy's dignity. *American Political Science Review* 106(4): 827–846.

Ober, Joshua. 2015. *The rise and fall of classical Greece*. New York: Penguin.

O'Brien, Patrick K. 2011. The nature and historical evolution of an exceptional fiscal state and its possible significance for the precocious commercialization and industrialization of the British economy from Cromwell to Nelson. *Economic History Review* 64: 408–446.

O'Donnell, Guillermo, and Philippe C. Schmitter. 1986. *Transitions from authoritarian rule*. Baltimore: Johns Hopkins University Press.

Organization of Economic Cooperation and Development. 2013. *OECD public governance reviews Colombia: Implementing good governance*. Paris: OECD.

Osafo-Kwaako, Philip, and James A. Robinson. 2013. Political centralization in pre-colonial Africa. *Journal of Comparative Economics* 41(1): 534–564.

Pincus, Steven C. A., and James A. Robinson. 2012. What really happened during the glorious revolution? In *Institutions, property rights and economic growth: The legacy of Douglass North*, ed. Sebastián Galiani, and Itai Sened. New York: Cambridge University Press.

Pinzón Quijano, Joaquín, ed. 1977. *Senado de la República. Historia de las Leyes. Legislatura de 1961. Reforma Social Agraria* vol. I. Bogotá: Imprenta Nacional.

Przeworski, Adam. 2009. Conquered or granted? A history of suffrage extensions. *British Journal of Political Science* 39(2): 291–321.

Purdekova, Andrea. 2011. 'Even if i am not here, there are so many Eyes': Surveillance and state reach in Rwanda. *Journal of Modern African Studies* 49(3): 475–497.

Putnam, Robert H. 1993. *Making democracy work*. Princeton: Princeton University Press.

Reyntjens, Filip. 2015. *Political governance in post-genocide Rwanda*. New York: Cambridge University Press.

Roberts, Michael. 1956. The Military Revolution, 1560–1660. Reprinted with some amendments in his *Essays in Swedish History*. London: Weidenfeld and Nicholson.

Robinson, James A. 2013. Colombia: Another 100 years of solitude? *Current History* 112(751): 43–48.

Robinson, James A. 2016. La Miseria en Colombia. *Desarrollo y Sociedad* 76(1): 1–70.

Ronderos, María Teresa. 2014. *Guerras Recicladas*. Bogotá: Aguilar.

Rothstein, Bo, and Dietlind Stolle. 2008. The state and social capital: An institutional theory of generalized trust. *Comparative Politics* 40(4): 441–459.

Rueschemeyer, Dietrich, Evelyn H. Stephens, and John D. Stephens. 1992. *Capitalist development and democracy*. Chicago: University of Chicago Press.

Sebarenzi, Joseph. 2009. *God sleeps in Rwanda*. New York: Atira Books.

Sahlins, Marshall D. 1961. The Segmentary lineage: An organization of predatory expansion. *American Anthropologist* 63: (2:1): 322–345.

Saylor, Ryan. 2014. *State building in boom times: Commodities and coalitions in Latin America and Africa*. New York: Oxford University Press.

Scott, James C. 1998. *Seeing like a state: How certain schemes to improve the human condition have failed*. New Haven: Yale University Press.

Scott, James C. 2010. *The art of not being governed*. New Haven: Yale University Press.

Skinner, Quentin. 1978. *The foundations of modern political thought*, 2 vols. New York: Cambridge University Press.

Skocpol, Theda. 2003. *Diminished democracy: From membership to management in American Civic Life*. Norman: University of Oklahoma Press.

Skowronek, Stephen. 1982. *Building a New American state: The expansion of national administrative capacities, 1877–1920*. New York: Cambridge University Press.

Slater, Daniel. 2010. *Ordering power: Contentious politics and authoritarian leviathans in Southeast Asia*. New York: Cambridge University Press.

Snodgrass, Anthony M. 1980. *Archaic Greece: The age of experiment*. Berkeley: University of California Press.

Soifer, Hillel. 2015. *Authority over distance: Institutions and long-run variation in state development in Latin America*. New York: Cambridge University Press.

Sommers, Marc. 2012. *Stuck. Rwandan youth and the struggle for adulthood*. Athens: University of Georgia Press.

Spruyt, Hendrik. 1994. *The sovereign state and its competitors*. Princeton: Princeton University Press.

Spruyt, Hendrik. 2009. War, trade and state formation. In *The Oxford Handbook of comparative politics*, ed. Carles Boix and Susan C. Stokes. New York: Oxford University Press.

Sreenivasan, Govind P. 2013. Beyond the village: Recent approaches to the social history of the early Modern German peasantry. *History Compass* 11(1): 47–64.

Strathern, Andrew. 1975. *The rope of Moka: Big-men and ceremonial exchange in Mount Hagen New Guinea*. New York: Cambridge University Press.

Strauss, Scott, and Lars Waldorf, eds. 2011. *Remaking Rwanda: State building and human rights after mass violence*. Madison: University of Wisconsin Press.

Therborn, Goran. 1977. The rule of capital and the rise of democracy. *New Left Review* CIII: 3–41.

Thies, Cameron G. 2007. The political economy of state building in Sub-Saharan Africa. *Journal of Politics* 69(3): 716–731.

Thies, Cameron G. 2005. War, rivalry, and state building in Latin America. *American Journal of Political Science* 49(3): 451–465.

Thompson, Edward P. 1971. The moral economy of the English crowd in the eighteenth century. *Past and Present* 50: 76–136.

Thompson, Edward P. 2001. *Customs in common: Studies in traditional popular culture*. London: The New Press.

Tilly, Charles, ed. 1975. *The formation of national states in Western Europe*. Princeton: Princeton University Press.

Tilly, Charles. 1978. *From mobilization to revolution*. New York: Addison-Wesley.

Tilly, Charles. 1990. *Coercion, capital and European states*. Oxford: Basil Blackwell.

Tilly, Charles. 1995. *Popular contention in Great Britain, 1758 to 1834*. London: Paradigm Publishers.

de Tocqueville, Alexis. 1856. *Ancien regime and the French revolution*. New York: Harper and Brothers Publishers.

de Tocqueville, Alexis. 2008. *Democracy in America*. New York: Penguin.

Turner, Frederick Jackson. 1921. *The frontier in American history*. New York: Holt.

Vansina, Jan. 2004. *Antecedents to modern Rwanda: The Nyiginya Kingdom*. Madison: University of Wisconsin Press.

Verwimp, Philip. 2013. *Peasants in power: The political economy of development and genocide in Rwanda*. Berlin: Springer.

Waldorf, Lars. 2007. Censorship and propaganda in post-genocide Rwanda. In *The media and the Rwandan genocide*, ed. Allan Thompson. London: Pluto Press.

Weber, Eugen. 1976. *Peasants into Frenchmen*. Stanford: Stanford University Press.

Wheeler, Nicholas C. 2011. The Noble enterprise of state building: Reconsidering the rise and fall of the modern state in Prussia and Poland. *Comparative Politics* 44(1): 21–38.

Wood, Andy. 2001. *Riot, rebellion and popular politics in early Modern England*. London: Palgrave Macmillan.

Wood, Andy. 2010. *The 1549 rebellions and the making of early Modern England*. New York: Cambridge University Press.

Wood, Andy. 2013. *The memory of the people*. New York: Cambridge University Press.

Wrightson, Keith. 1982. *English society, 1580–1680*. New Brunswick: Rutgers University Press.

Yamigizawa-Drott, David. 2014. Propaganda and conflict: Evidence from the Rwandan genocide. *Quarterly Journal of Economics* 129(4): 1947–1994.

States and Development: Early Modern India, China, and the Great Divergence

Bishnupriya Gupta, Debin Ma, and Tirthankar Roy

Abstract Can differences in state capacity explain the Great Divergence between Asia and Europe? Evidence from India and China suggests that customary property rights provided de facto rights to land and community ties substituted for de jure property rights in mercantile activities. Economic activity did not face an undue risk of expropriation. China and India generated lower fiscal revenue per capita compared to Europe. We explain the big difference in revenue per capita between the two Asian countries and England in the early modern period. In terms of the differences in the threat of internal and external conflicts. The large empires in Asia faced a disproportionate threat of internal rebellions and traded off fiscal capacity for appeasement of local ruling groups and their military support in external conflicts.

Keywords Property rights • Fiscal capacity • Warfare • Great divergence • Mughal India • Qing China • East India Company

Did states in different world regions promote or inhibit capitalist economic growth in the early modern times? This is now the subject of a large scholarship. The comparative history evidence tests whether world regions were similar or different in the eighteenth century, whether "modern" polities emerged in these regions at the same time or not, suggesting new theories and clarifying older ones about how state formation and political modernization make a difference to economic change. The scholarship on Europe is extensive and has recently made a big move into comparing Britain with China. India is yet to be firmly integrated in this comparative history of states. The present chapter is an attempt to apply hypotheses drawn from the Eurasia literature to early modern India and draw comparisons with China.

B. Gupta (✉)
University of Warwick, Coventry, UK
e-mail: b.gupta@warwick.ac.uk

D. Ma • T. Roy
London School of Economics, London, UK

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_2

The state provides the necessary conditions to achieve long-run economic growth. It does it in two ways mainly: supply of growth-promoting institutions such as protection of private enterprise and supply of public goods such as infrastructure and defense. The first refers to the rule of law, that is, a legal system that enforces property rights (North 1982; North and Weingast 1989). It can also be based on community enforcement mechanisms that allow trade and commerce and even industrial activity to grow (Greif 2006). The second way requires the state to grow in size and capacity to invest, and in turn, grow in military capability as well. The states that emerged in Asia, Africa, and Latin America before European expansion and colonization pursued these objectives in a different way from the states in Europe. These considerations give rise to two broad questions about the comparative history of states and their divergent economic outcomes. How was secure property defined and protected by the state? What were the conditions for sustainable growth in size and capability?

The present chapter explores the evidence from India and China in relation to Europe in order to shed some light on these questions. We start with a discussion of the ways in which the states in India and China differed from Europe and pursued the objectives of limiting risk of expropriation and providing security of state boundaries in the early modern period. We review the comparative history evidence by bunching these questions into two sets, one dealing with property right protection, and the other with state capacity.

Property Rights: A Comparative History Perspective

Douglass North's interpretive history of Western Europe emphasizes security of private property and the related idea that state power to expropriate needs to be constrained. An important critique of the case for limited government is provided by S.R. Epstein (2000), suggesting that jurisdictional fragmentation rather than constrained government was the key to early modern economic growth. Controversies over the European experience aside, the perspective raises the question, how secure were private property rights from expropriation in other regions of the world?

The question of manipulation of property rights arises quite strongly in the particular context of European expansion in the New World. Organized land-grab and labor servitude defined the features of what has been called "extractive" settler policy (Acemoglu et al. 2001). The evidence concentrates especially on land and natural resource extractive activities. As for the Old World regions like India or China, Karl Marx believed that private property did not exist at all in Asia (Thorner 1966). Eric Jones (1988) believes that private property was insecure in the East. Most definitions of property rights hinge on the feature of marketability of land. If customary rights on land make the cultivator liable for taxation, then the absence of the right to buy and sell land is, at best, a narrow definition of property rights.

The oriental despotism scholarship inspired by Marx considered that the emperors also oppressed capitalists in Asia, whereas their power was constrained in Europe. "In general", J. Barrington Moore wrote, "the attitude of the political authorities in India towards the merchants seems to have been closer to that of a spider towards a fly than that of the cowherd towards his cow that was widespread in Europe at the same time."[1] William Moreland (1923), the economic historian of medieval India, drew a picture of the Mughal state as an extractive machine, leaving little space for a commercial culture to develop except in elite luxury goods or in moving grain. The Marxist anthropologist Eric Wolf (1982) saw the effect of the "tributary" Mughal system upon merchants in much the same light.

Property Rights: Mughal and Post-Mughal India

Was property secure and widely accessible in Mughal India? The existing evidence suggests that pre-British states in India took private property and the security of property quite seriously, so that there was little the most important post-Mughal regime, the British imperial state, could add to this dimension. The Mughals maintained detailed revenue records, conducted cadastral surveys, and offered incentives to established settlers. These measures suggest that the Mughals recognized and valued private property. Confiscation of property was a conditionality of tax farming arrangements common in the Islamic empires of Asia. But confiscation did not mean taking over ownership rights but the withdrawal of an entitlement. An entitlement to enjoy the returns from property did not mean the same thing as owning the property.

From the late eighteenth century, much of India became a territory ruled by the East India Company. When in the late eighteenth century, the Company set out to create a new framework of law for territories it then controlled, they discovered more apparent similarities than differences between precolonial India and preindustrial Europe in the structure of economic laws. For example, the British discovered by perusing Sanskrit, Persian, and Arabic texts in the eighteenth century that the notions of legality and rule of law were as developed in India as in Europe. In some of these texts that they translated for use in the courtroom, the meaning of private property and the theory of the sanctity of contract closely resembled their European counterparts.

This does not mean that property rights were of the same kind in India and Europe. The peasant's property right was a right to use agricultural land but not exactly a proprietary right in the modern sense. It was not because the tax farmer or the military-fiscal tenure holder also had a right to use the same agricultural land in

[1] J. Barrington Moore, *Social origins of dictatorship and democracy*, 1967, p. 322.

a different way, but land could not be bought or sold without there being a political bargain of some kind. But that might not have adversely affected the peasant's incentive to improve land. The colonial government under the East India Company introduced a new land tenure system in India at the end of the eighteenth century to show that private ownership of land could be defined legally. The construction of property rights in land as a formal institution was new and de jure and gave landowners the right to buy and sell. But the customary rights before the arrival of the British carried with it "informal" rights to this land defined by the responsibility to pay taxes and intergenerational rights to cultivation.

By 1860, India was the largest colony in the British Empire. European settlers were few in number in the region, although they did exercise a great deal of influence on British imperial and foreign policy in the nineteenth century. Did the settlers also play a role in the allocation of resources? The evidence that they did so in an institutional way was largely absent in India. In land rights, British Indian legislation did not discriminate between people by the ethnicity of the right-holder. If anything, expatriate land-holding rights were weak compared with indigenous rights until well into the nineteenth century. Europeans could not own or purchase farmlands, for example, until the late 1830s, 80 years after colonization had begun. Agricultural land was owned by peasants and landlords, and the British legislated on these rights so heavily that any land transfer, because it involved a complicated legal process, became progressively rare. Only in the case of tea and coffee cultivation were forest lands leased out by the Government to plantation companies, but these actions did not involve transfer of property rights held by Indians before.

Did the Mughal state expropriate private capitalist property routinely? Recent rethinking of the Mughal state has more or less discarded the Marx–Moreland approach suggesting a huge imbalance in the relation between the state and the capitalist order, emphasizing instead its dependence on mercantile classes for a smooth running of the fiscal and monetary systems. While bankers could be pressurized to lend to the courts, there is little evidence that their property was under threat of usurpation in the Indian cities, barring sedition or disloyalty. Private trade developed and international trade saw rapid growth in the seventeenth and eighteenth centuries with the entry of the European trading companies. Trading rights were granted to these companies by the monarch and the risk of expropriation did not hinder India's domination of the international textile market. The rise and the expansion of textile trade and other market exchanges used community-based rules to provide a secure institutional setting for commercial activities.

Insurance rates on trade, which would reflect the risk of expropriation, were not abnormally high. Community enforcement mechanisms prevailed in the communities of Indian traders and bankers. The absence of such contract enforcement mechanism between the European companies and Indian producers and traders was seen as a shortcoming of the legal system for outsiders to the customary laws. For the Indian merchants customary laws were adequate.

Property Rights: China[2]

China's Tang rulers shifted the main source of tax revenue from labor to land at the end of the eighth century. From the perspective of tax collection, this put a premium on establishing clearly defined property rights in land. Accordingly, the state began to relinquish control and regulation of land tenure. While the emperor retained de jure control over all land, this new arrangement contributed to the emergence of de facto private land ownership and fostered the rise of a family-based owner-cum-tenant system of agricultural cultivation that dominated the economy of imperial China for the next millennium (Ma 2012).

The emerging private property rights over land typically included residual claimancy; the right to rent, sell, or mortgage land; and the right to bequest. Recognition of these rights allowed private owners rather than the state to capture the benefits associated with rising productivity, new land reclamation, and population growth. If households became unable to farm the land themselves, the opportunity to transfer their land-use rights via sale or rental allowed them to capture the returns to their investments in the land.

Although de facto private property rights in land date back to much earlier periods, the Tang reforms encouraged the development of increasingly sophisticated markets for land in which layers of ownership and user rights could be purchased, sold, rented, mortgaged, and divided.[3] Ownership of a single plot could be vested in separate parties endowed with rights over the surface and sub-surface respectively – the so-called *yitian liangzhu* (two lords to a field) or even *yitian sanzhu* (three lords to a field) system – rights that could then be sold, leased, or used as collateral. Tenants as well as owners could freely buy, sell, lease, or mortgage their access rights. The efforts of millions of households to use land as a vehicle for maintaining and, if possible, improving their socioeconomic status produced a variety of complex outcomes. Given the irregular income stream associated with farming, the rich array of vehicles that mimicked many contemporary financial instruments allowed average farmers to "financialize" land, their premier store of wealth, to provide liquidity and stabilize their income. The same features also facilitated payment of land taxes.

Revenue constraints at all levels of government led to a common pattern: officials granted merchants, guilds, or kinship groups unwritten but well understood and enforceable commercial property rights in return for a stream of tax revenues, in what amounted to a form of tax-farming. These accommodations typically involved agreements between officials and commercial groups (rather than individual merchants). Mercantile groups selected leaders whose authority was recognized by both officials and fellow merchants. Officials expected these "head merchants" to deliver

[2]The narrative on this section draws heavily from Brandt et al. (2014).

[3]Land transactions were often recorded in written documents, many of which survive in libraries and archives. See citations in Brandt et al. (2014).

tax revenues and control the actions of their associates. In return, officials stood ready to utilize their power—either directly or by allowing authorized merchants or their agents to act as informal official deputies—to block initiatives that threatened the operations of the "legitimate" or "insider" merchants (and hence the established revenue streams). Ho Ping-ti (1954) observes that "the interventionist state … protected the vested interests of all salt merchants".[4]

Such structures equipped mercantile groups with substantial control of their trades, at least within the territories administered by cooperating officials. Management of specific trades was often farmed out in this fashion to mercantile guilds, which typically allowed newcomers to enter their business provided that they obeyed guild rules relating to currency, weights and measures, product quality, apprenticeship, wages, piece rates, and, of course, tax payments.

These practices, which essentially amounted to official sales of market control in exchange for informal revenues, had important implications. The presence of entrenched commercial interests with backing from local or even central officials may have acted as a long-term brake on innovation. The prevalence of irregular taxation at all levels of government helps to explain the apparent contradiction between the low rates reflected in the receipts of the Board of Revenue (Table 2) and the popular image of Ming and Qing as rapacious regimes (Ma 2014).

The lack of legal formalization in commercial and civil matters introduced an element of uncertainty into private ownership. Although private property rights in Ming-Qing China were genuine and substantive, from the official perspective, private ownership remained secondary or derivative from the political standing of the property holders. Faced with the possibility, however remote, of confiscatory intervention that no legal response could remedy, property holders felt it necessary to seek shelter under an umbrella of political power (Deng Jianpeng 2006, p. 69). Despite elaborate and generally predictable informal arrangements for recording, protecting, and transferring rights over land and other tangible assets, the foundation of property rights in imperial China, particularly in the sphere of commerce, rested on politics rather than law, with implications for economic change during the nineteenth century and beyond, to be discussed later.

Although these property rights may not be real serious impediments to agricultural, and household production, or even informal and small-scale markets, these rights (and the underlying legal system) seem largely responsible for the restricted development of financial instruments, particular the absence of any formal market instruments for governmental or public debt, an area in which China looks much different from pre-industrial European nations, particularly Holland and England. De facto governmental borrowing (or extraction) existed in the form of advanced collection of taxes, forced "borrowing" from merchants, sale of official titles, and so on certainly existed. Any form of formal and institutionalized public debt instruments and markets would be hard to sustain in an environment where imperial power trumped whatever little bargaining power the merchant financiers

[4] Ho, Ping-ti (1954). p. 142.

may have had in traditional China. Indeed, a more formalized imperial or regional governmental borrowing in China started only in the nineteenth to twentieth-century treaty port era when the bargaining power of merchants and bankers (particular Westerners) were backed up by extraterritorial privileges or Western consular or military presence. The absence of orderly and credible financial instruments such as public debt restrained state capacity and rendered the traditional Chinese state prone to outright fiscal predation or confiscation in times of political crises or widespread breakdown in social order (Ma 2014).

State Capacity: The European and World History Perspective

Was the state capable? Whatever the status of private property, states matter to economic development because they supply public goods such as law, policing, and defense. These activities cost money, and therefore a state with a higher capacity to extract resources can in theory offer more of these goods. A simple measure of state capacity is fiscal capacity, tax per person, or tax per square mile of the territory under control; and comparative data on this measure shows wide differences between seventeenth-century western European states and Asiatic states (Table 1). The role of state capacity on economic development has been demonstrated in a number of recent works. Karaman and Pamuk (2010) have shown this for the Ottoman Empire and the Mughal Empire. Ma (2014) makes a similar point for the Qing Empire, and Vries (2013) demonstrates China–Britain differences with earning and spending data.

Fiscal capacity depends on the bureaucratic ability of the state to collect revenue. The ability of the taxpayer to pay also matters in terms of the volume of revenue. Richer societies can pay more taxes. In poorer societies, where a majority of the population is at a subsistence level, the capacity of the state to tax is also limited. The relationship between per capita taxation and per capita income is nonlinear. We show per capita tax rates declined in India and China until 1800. This is consistent with declining per capita GDP for India and China over the Long Eighteenth Century (Bassino et al. 2015). Economic development is endogenous to fiscal capacity. To determine a causal relationship, we need an exogenous factor. In the European literature these are wars. Post-Black-Death Britain never went back to a

Table 1 Tax revenue per head in grams of silver

	India	China	Ottoman Empire	France	Spain	England	Dutch Republic
c. 1670	20	7	12	46	36	45	
c. 1800	15	3–4	13–15	66	63	160–303	171
c. 1850	16	7				344	

Source: The figures for India have been calculated from the revenue data in Table 2 in Roy (2013). The figures for China are from Ma (2014). For other countries the data are from Karaman and Pamuk (2010), available at http://www.ata.boun.edu.tr/sevketpamuk/JEH2010articledatabase

Malthusian level due to the high mortality regime of wars and urban disamenity (Voigtländer and Voth 2013). The effect of military conflict in the years before 1800 operated through mortality shocks that raised per capita GDP and the capacity to pay taxes. Such instruments were missing in India and China.

What caused big differences in the capacity to extract resources? On this question, Vries and others rely on a factor, which Max Weber once thought was an important characteristic of modern or rational–legal states: a bureaucracy that is professional and acts without conflict of interest. State capacity can be constrained also by geographical factors, such as low productivity of land, and high trade costs, making it more difficult for states to raise resources by taxing trade. In criticisms of the settler model, Jeffrey Sachs (2012) has argued that a preoccupation with settler policy carries an oversight of geographical differences between tropical and temperate regions. Roman Studer has shown persuasively that costs of market exchange varied a great deal between early modern Europe and India due to geographical factors. Moving goods from the interior to the ports was a far more expensive operation in India in the 1700s than it was in Western Europe. States, in that setup, were constrained by the high cost of trade (Studer 2015).

A related question is: Was the state strong? States should be not only large but also capable of withstanding threats of invasion and rebellion in order to undertake investments. A subtheme in the Europeanist literature on the enlargement of the state in early modern times is the growth of military capacity. Some of the major European states in the eighteenth century moved towards sovereign control of the fiscal and the military organizations, away from dependence on mercenaries, creditors, and contractors. The phrase "military fiscalism" is used to suggest a co-evolution of fiscal capacity and military capacity (Wolfe 1972). The corresponding growth of "social power" through a process of conflict brought about the nation state and the state structures that defined European modernization (Mann 1986).

A great deal of the recent discussion on the comparative history of state capacity has focused on warfare. Following Charles Tilly (1989), many have argued that military conflicts created state capacity.[5] The European states introduced mechanisms to raise fiscal capacity in order to finance wars. Tilly argued that frequent wars in Europe created the need to finance militarization and therefore saw a rise in fiscal capacity. Hoffman (2011) distinguishes between the period before and after 1800, when military defeat under Napoleon led to loss of power. When monarchs were threatened with a loss of power, they invested in building military capability and introduced new administrative controls. Gennaioli and Voth (2015) and Dincecco and Prado (2012) find a positive correlation between wars and fiscal capacity. Does this have universal validity? Besley and Persson (2008) show that

[5]Mobilization of resources for war as a catalyst in the making of a modern fiscal system is emphasized in O'Brien 1988; and in the formation of nation states by B. Downing (1991). The military revolution and political change in early modern Europe. For an economic analysis of the contribution of war to state capacity, see Besley and Persson (2009). A recent paper argues that political competition induced investment in military technology, and the beginnings of a "military revolution" in England, France, and Germany, Hoffman 2011.

internal conflicts reduced state capacity, whereas external conflicts increased it. Besley and Persson (2009) show a positive correlation between fiscal capacity, legal capacity, and warfare.

The standard European account of warfare and state development predicts that historical warfare produced greater fiscal capacity and long-run economic development. These were mainly external conflicts. Dincecco et al. (2016) show that this argument does not hold for low population density African states with very different labor market institutions that generated conflict over labor use rather than territory. Greater prevalence of internal conflicts may lead to different outcomes too. Large empires such as China and India may have faced two different options: Appeasement and repression. The monarch may co-opt potential challengers through threat of repression or create an alliance with them through lower taxation when faced with the threat of a military challenge or rebellion With a multitude of ruling elites these empires therefore generate lower levels of revenue at the center. This literature raises two questions. Why was state capacity relatively small in Asia compared with Europe? (Table 1). And can the Asian divergence be assessed in terms of warfare?

State Capacity: The Indian Evidence

It is well established now that early modern and modern states in Asia were much smaller compared with contemporary European states in fiscal capacity. And yet, the Ottoman, the Mughal, the Qing were empires that lasted a long time. Surely weak fiscal capacity did not mean military weakness. What then did military power derive from?

Scholars of Europe have too readily wedded themselves to the idea, which Max Weber first introduced in his analysis of the modern state, that a state is an entity that enjoys a monopoly of violence.[6] Mancur Olson (2000) gave the idea a historical twist by suggesting that states evolved in the manner of a bandit that gives up oppressing its flock in unpredictable fashion and begins to collect predictable protection money (tax) from them instead. Both ideas, now famous clichés in economic history, suggest that a centralization of military power is a feature of any powerful state. The idea obscures rather than illuminates the character of Asian empires. There was no such thing as a centralization of power in Turkey, India, or Persia in the seventeenth century. Although China had the most centralized and hierarchical system, the size of the center, as we will see later, remained fiscally modest and militarily limited. The fiscally weak entities were heavily dependent on local constituents who were allowed to earn money on the side. And yet the imperial center did command military power when under serious threat. The West and South Asians achieved this by means of a cooperative arrangement with semi-independent

[6] For a discussion, see Berman 1987.

local chiefs and military elite. Sharing sovereign powers rather than concentrating it at the top was the key to their survival. The offer of temporary revenue grants or *iqta*[7] was an instrument of securing cooperation, and the withdrawal of that offer at times signified that it was a conditional entitlement rather than a property right.

Recent interpretations of the Mughal state and fiscal system have revised earlier views of a highly centralized state, by placing a great deal of stress on the dependence of the political center upon local power structures, consisting of the landlords and military tenure holders known as *jagirdars* and *zamindars*, among others (Hasan 2004). The concept of a decentralized polity is found to be a useful construct also for medieval Deccan and Maratha states and for Northern India itself (Stein 1985; Perlin 1985; Wink 1986; Ali 1993). Kathleen Gough (1979 and 1981) interpreted the Chola state in South India as a regime that allowed communities to retain substantial social and economic autonomy. Another closely related notion is suzerainty, wherein taxpayers enjoy some domestic autonomy in exchange of a promise to pay taxes, as opposed to sovereignty. Suzerainty was applied in the context of the Ottoman imperial state and found useful by Aidan Southall (1988) to describe the Chola and Rajput lineage states. The principle of leaving lineages alone was common between those examples of Africa and South Asia, in which the term "segmentary state" was used.

The military strength of such a regime flows from bottom to the top, from the army units contributed by local chiefs and communities. In principle the local chiefs would weigh the advantage of staying in the coalition as a security against rebellion or invasion in their own domain against the disadvantage of having to pay regular tribute to the center. If, perceiving that the disadvantage weighed more than the advantage, an individual chief rebelled, the risk of a crushing retribution was too great. If a number of them reasoned in the same way and rebelled together, the empire would collapse. The Mughal emperor tried to protect his interest by maintaining a sound system of intelligence and by the bifurcation of the governance of provinces into two heads, the head of civil affairs and the head of military affairs, the latter being usually recruited from immediate family and the former from among competent court officers. At the turn of the eighteenth century, nevertheless, the balance did seem to change as a simultaneous outbreak of rebellions rendered the center suddenly vulnerable. In a remarkably short period of time, the Mughal Empire lost major provinces to nominally loyal but fiscally independent governors, and a number of border zones to rebels and enemies.

The collapse of the coalition turned the members of the coalition upon each other. The dominant tendency in the second half of the eighteenth century was an increasingly violent contest for revenue between regimes that had succeeded the Mughal Empire. The breakup of the empire, and the struggle for revenue among contenders, unleashed much potential for violence. Rulers, noblemen, warlords, and underneath them, landlords and peasant clans more or less lived in a state of war, especially in western and northern India. All of them sought to improve the means

[7] Land grant given to a local feudal lord.

to acquire more money to finance warfare. But the outcome of the contest was not collective empowerment of the states after the pattern of Western Europe, but quite the opposite, a collective disempowerment. In keeping up the military enterprise, almost all of these political actors ran into fiscal crisis and the states or quasi-states that they had formed shrank in size.

There was one large exception to this picture, and that was the British East India Company. The Company gained in military power through the conflict. Did it gain also in fiscal power? It did not. There is no evidence that the overall fiscal capacity of the state had dramatically improved after British takeover. The Company did, however, manage to reduce the dependence of the government upon local powers. This required a twofold change. First, converting entitlements to property into ownership right, which reduced the power and the military role of the former holders of military tenures. Second, creating a standing army. There was a redistribution of sovereign power and military power at the same time. The process was far from a straightforward one, but it did deliver a degree of centralization that did not exist in Indian history before, at least not on such a scale as in the nineteenth century. The Company's strategy to concentrate military and fiscal power at the center made it a distinctively new kind of rule in India, and an essentially European one, in comparison with most of its Indian contenders. New scholarship shows why the Company succeeded as a military-fiscal state whereas its Indian rivals often failed, and emphasizes the European dimension in the Company's strategy and statecraft in late-eighteenth-century India (Oak and Swamy 2012; Roy 2013).

The centralization of the army was reinforced by the Company's already established position as a naval power. Whereas most empires of the past had a tenuous access to the seaboard, the Raj controlled the seaboard firmly and used that control to foster maritime trade. The Raj emerged not from outright conquest but from the activities of the East India Company between 1765 and 1818. The Company was a merchant firm. Although it had left its commercial legacy far behind in 1858 when it was removed from the formal rule of India, it understood the commercial importance of India, especially its port cities.

The centralization of military power made the East India Company and the Crown rule that followed it, a strong state, and a European experiment in India, but did not make it a more capable state in terms of spending power. The Raj ran a small government. In the 1920s, nominal tax collection per capita was 6 % of tax per head in Britain adjusted by purchasing power parity (Roy 2012). British India was poor also in relation to most of Britain's tropical colonies and other Asian countries in the interwar period. Between 1920 and 1930, the government of the Federated Malay States spent on average more than ten times the money spent in British India per head, that of Ceylon spent more than three times, those of the Philippines and the Dutch East Indies more than double, and those of Siam and French Indochina 40–50 % more (Roy 1996). The Raj was a small government also in relation to its own national income. Government revenue as a proportion of national income was 2 % in 1871 and marginally higher at 3–5 % in 1920–1930. It was 19 % in Britain and 29 % in Japan in the interwar years (Roy 2012).

The result of this fundamentally European experiment in India was distinctly modern if we measure modernity by military capacity. For the first time in Indian history a state did enjoy an absolute monopoly in the exercise of violence. But it was a distinctly traditional Indian indigenous state if we measure modernity in terms of spending power, especially spending on public goods. Barring canals in some regions, and limited road-building projects, public goods such as health and education were created only on a modest scale. The cities were better provided for than the countryside. Growth of trade and the manufacturing industry created a demand for higher education, spurring private investment as well as provincial government investment in education, but again, the scale of spending was biased towards the richer regions and the port cities where most of the wealthy capitalists lived.

State Capacity: The Evidence on China

The biggest threats to the Chinese emperor, with absolute hereditary power and without formal constraint on its rule, came from external invasion or internal insurrection. Rebellions were frequent throughout Chinese history. While external security required additional revenues, internal rebellion required distribution of taxation to ensure a satisfactory level of income for the peasantry. China's fiscal system was centered on the taxation of privately owned land.[8] For simplicity, we assume a fixed stock of taxable land, so that increasing fiscal revenue depended on the official tax rate on a unit of land and the expenditure on the bureaucracy devoted to tax collection. Government revenue was increasing in both.

In this very simple setting, the problem facing the Emperor was to determine the tax on land and the size of the bureaucracy that would maximize net fiscal revenue (total tax collection less administrative costs) and at the same time prevent discontent among farm households by leaving them with a decent post-tax income. The land-tax system created sharp conflicts between the tax collectors and land-owning households, including local elites, who often held substantial acreage. There was unofficial tax farming that allowed the local officials to tax the peasantry for their private benefit. The peasants resisted such impositions through violent protests and allied themselves with the local gentry. High agency costs associated with administration of the revenue system and concerns about insurrection steered imperial China towards modest formal taxation, a small official bureaucracy, and a correspondingly limited scope of non-military activities financed from the public purse.

[8] Yeh-chien Wang (1973) shows that the share of land tax in total fiscal revenue was 74 % in 1753 (p. 80). Other sources of taxes included revenue from state monopolies, particularly on salt. The predominant share of land tax in total taxation can be observed throughout the Ming and Qing dynasties. Ibid.

Following the uprisings during the early years of the Qing dynasty as order was re-established, formal tax revenue expressed in silver taels remained fairly constant between 1700 and 1850, fluctuating around an annual average of 36 million silver taels, of which approximately 70 % came from taxes on land.[9] With stable revenue and substantial growth of population, per capita tax revenue fell steadily: by 1850, per capita revenue was less than half the level for 1700. Nominal revenues rose sharply in the late nineteenth century, but the increase was modest in real terms. The share of government revenue in total output during Qing was low: Yeh-chien Wang finds that late-nineteenth-century land taxes represented 2–4 % of the produce of the land in most areas, although it may have been a larger share in the prosperous Yangzi delta region, and that total government revenue from all sources amounted to roughly 2.4 % of net national product in 1908 (Ma 2014; Wang 1973).

Comparative data shown in Table 1 demonstrate the Qing dynasty's limited fiscal capacity. During the late eighteenth century, per capita revenues were the lowest in China not only in comparison with the European states, but also in comparison with India and the Ottoman Empire. England's government revenue even surpassed the comparable figure for the immensely larger and more populous Qing Empire! The data also show that per capita revenues, which remained roughly constant over long periods under the Qing, tended to increase in most countries (See Fig. 1). Expressing the tax burden in terms of the number of days an unskilled urban worker would have to work in order to earn the equivalent of the average individual tax payment provides an alternative view of the Qing Empire's modest fiscal capacity (Table 2).

Limited revenues meant that the size of the bureaucracy lagged far behind the growth of population. Indeed, the number of counties (*xian*), units ruled by magistrates who occupied the lowest rung of the official hierarchy, hardly changed after Han times (206 BC–AD 220). Despite a vastly larger population and territory, Qing China had only 1360 counties compared to 1230 under the Song (Skinner 1977). Limitations on the size of the bureaucracy help explain why the imperial administration never penetrated below the county level. With a near-static administrative structure, population growth meant that an average Qing county had more than triple the inhabitants of an average Song county. This increased the administrative burden facing local magistrates and magnified their dependence on the active cooperation of local gentry, which in turn reinforced the pressures mandating a low-tax regime, as efforts to increase taxes would place local officials in direct conflict with the economic interests of the same elites whose advice and cooperation was needed to manage local affairs and preserve social order.

Like all revenue-constrained pre-modern empires, Qing's public spending was limited to programs of external and internal security. A large empire like China with

[9] One Chinese silver tael = 37 g of silver.

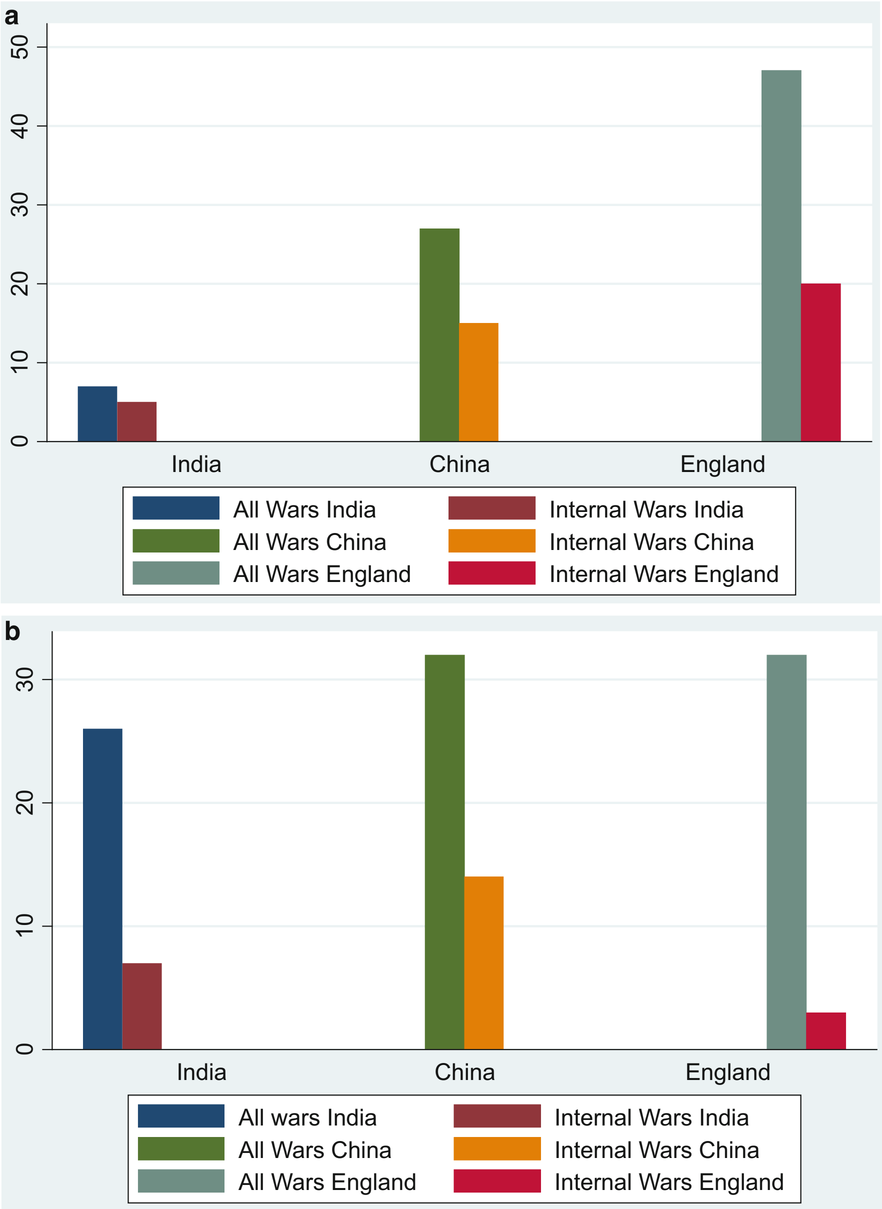


Fig. 1 (**a**) Number of wars in the seventeenth century. (**b**) Number of wars in the eighteenth century. (**c**) Number of wars in the nineteenth century (Source: Brecke Data Set, Conflict Catalog (Violent Conflicts 1400 A.D. to the Present in Different Regions of the World), Centre for Global Economic History, Utrecht)

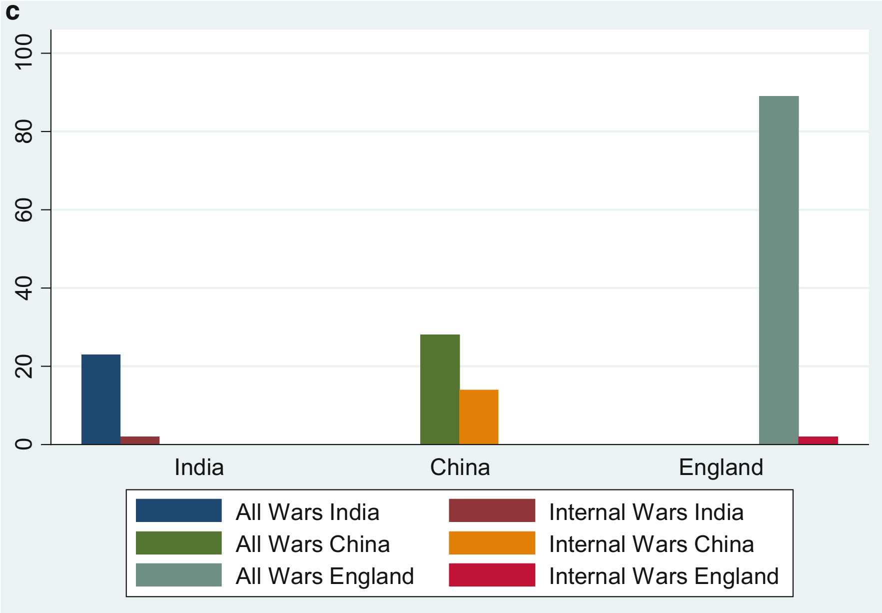


Fig. 1 (continued)

Table 2 International comparison of per capita tax revenue in days of urban unskilled wages

	China	Ottoman	Russia	France	Spain	England	Dutch R
1650–1699		1.7		8.0	7.7	4.2	13.6
1700–1749	2.26	2.6	6.4	6.7	4.6	8.9	24.1
1750–1799	1.32	2.0	8.3	11.4	10.0	12.6	22.8
1800–1849	1.23					17.2	
1850–1899	1.99					19.4	

Source: See Ma (2014). The figures for China are from Ma (2014). For other countries the data are from Karaman and Pamuk (2010), available at http://www.ata.boun.edu.tr/sevketpamuk/JEH2010articledatabase

limited revenues and a small bureaucracy was therefore constrained to spend on public goods. One area where they focused was maintaining increased agricultural productivity – for example investments in water management and irrigation – and on the operation of official granaries that could stave off famine and hence maintain public order when natural disaster struck. Such programs contributed directly to mitigate public discontent and thus supported the longevity of the imperial system (Brandt et al. 2014).

Warfare and State Capacity: An India–China Comparison

Can the pattern of warfare in India and China explain why a divergence emerged in fiscal capacity? If the frequency of wars benefited Europe by raising fiscal capacity after 1800, what was the evidence on warfare in China and India?

Figure 1 shows a comparison of warfare in India, China, and England. We consider the seventeenth, eighteenth, and nineteenth centuries and measure the number of wars in each year summed up to the total in each century.[10] We find that China experienced a large number of conflicts in the seventeenth century. However, an overwhelming majority were internal conflicts. This pattern continued in the eighteenth century. The very low revenue per capita in China (Tables 1 and 2) may well suggest a policy of appeasement as suggested in our discussion. In contrast most of the wars in England were external conflicts. The number of such conflicts in England rose sharply in the eighteenth and nineteenth centuries with colonization of the Americas, Asia, and Africa.

India shows a relatively low incidence of both external and internal conflicts in the seventeenth century and a higher revenue per capita compared to China (Table 1). External conflicts rose from the eighteenth century as European powers began their quest for colonization. Many of England's external conflicts from the late eighteenth century were imperialist wars in different parts of the world rather than only in Europe. India's external conflicts in the eighteenth and nineteenth centuries were with European powers first seeking new markets in India and then acquiring territorial rights. By the mid-nineteenth century, most of India was under British rule.

If wars are conflicts over markets, then large empires with large markets may be less prone to external conflicts. As in Besley and Persson (2008), external conflicts raised fiscal capacity. For large empires internal conflicts were the real threat. A monarch with long-term interest in governing tended to keep taxation at a level that prevented the threat of a rebellion. Evidence from India and China suggest that monarchs used various mechanisms to deal with the threat of internal rebellions. This equilibrium sustained the tenure of the Mughal Empire until the middle of the eighteenth century. Chinese empires similarly managed the threat of rebellions through various concessions. As suggested by Besley and Persson (2008), the internal conflicts reduced fiscal capacity rather than enhance it. Large Empires developed alternative means of militarization through decentralized support for local rulers and tax collectors. Therefore warfare rather than increasing fiscal capacity may have lowered the incentive to tax the producers and also prevent accumulating revenues centrally. Decentralization of revenue and its use in appeasement worked adversely for building fiscal capacity in large empires.

[10]We can also show the number of war years in each century, which takes into account the length of each war. The picture does not change.

Conclusion

Can the origin of the Great Divergence be explained in terms of the quality of state capacity? This chapter argued on the basis of evidence from India and China that we have to be careful in applying concepts derived from European history to these large Asian regions. In both India and China customary property rights worked as effective mechanisms against risks of expropriation and allowed economic activity to flourish. State capacity provides a more promising account of divergence. But the underlying causes of difference in state capacity need to be interpreted cautiously. This important qualification applies especially to the connection drawn between warfare and state formation in recent world history scholarship.

References

Acemoglu, Daron, Johnson Simon, and James Robinson. 2001. Colonial origins of comparative development: An empirical investigation. *American Economic Review* 91: 1369–1401.

Ali, M. Athar. 1993. The Mughal polity – A critique of revisionist approaches. *Modern Asian Studies* 27(4): 699–710.

Bassino, Jean Pascal, Stephen N. Broadberry, Kyoji Fukao, Bishnupriya Gupta and Masanori Takashima. 2015. Japan and the Great Divergence. 725–1874, mimeo.

Berman, H.J. 1987. Some false premises of Max Weber's sociology of law. *Washington University Law Review* 65: 758–770.

Besley, Timothy, and Torsten Persson. 2008. Wars and state capacity. *Journal of the European Economic Association* 6(2–3): 522–530.

Besley, Timothy, and Torsten Persson. 2009. The origins of state capacity: Property rights, taxation, and politics. *American Economic Review* 99(4): 1218–1244.

Brandt, Loren, Debin Ma, and Thomas G. Rawski. 2014. From divergence to convergence: Re-evaluating the history behind China's economic boom. *Journal of Economic Literature* 52(1): 45–123.

Deng, Jianpeng. 2006. *Caichan quanli de pinkun* [The poverty of property]. Beijing: Law Press.

Dincecco, Mark, and Prado Mauricio. 2012. Warfare, fiscal capacity, and performance. *Journal of Economic Growth* 17(3): 171–203.

Dincecco, Mark, James Fenske and Onorato Massimiliano. 2016. *Is Africa different? Historical conflict and state development*. Mimeo.

Downing, B. 1991. *The military revolution and political change in early modern Europe*. Princeton: Princeton University Press.

Epstein, Stephan R. 2000. *Freedom and growth: The rise of states and markets in Europe, 1300–1750*. London: Routledge.

Gough, Kathleen. 1979. Dravidian Kinship and modes of production. *Contributions to Indian Sociology* 13(2): 264–292.

Gough, Kathleen. 1981. *Rural society in Southeast India*. Cambridge: Cambridge University Press.

Greif, Avner. 2006. *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge: Cambridge University Press.

Hasan, Farhat. 2004. *State and locality in Mughal India: Power relations in Western India, C.1572–1730*. Cambridge: Cambridge University Press.

Ho, Ping-ti. 1954. The salt merchants of Yang-Chou: A study of commercial capitalism in eighteenth-century China. *Harvard Journal of Asiatic Studies* 17(1/2): 130–168.

Hoffman, Philip T. 2011. Prices, the military revolution, and Western Europe's comparative advantage in violence. *Economic History Review* 64(s1): 39–59.
Jones, Eric. 1988. *Growth recurring: Economic change in world history*. New York: Oxford University Press.
Karaman, K.K., and Şevket Pamuk. 2010. Ottoman state finances in European perspective, 1500–1914. *The Journal of Economic History* 70(03): 593–629.
Ma, Debin. 2012. Political institution and long-run economic trajectory: Some lessons from two millennia of Chinese civilization. In *Institutions and comparative economic development*, ed. Masahiko Aoki, Timur Kuran, and Gerard Roland, 78–98. Basingstoke: Palgrave Macmillan.
Ma, Debin. 2014. State capacity and great divergence, the case of Qing China (1644–1911). *Eurasian Geography and Economics* 54(5–6): 484–499.
Mann, Michael. 1986. *Sources of social power*, vol. 1. Cambridge: Cambridge University Press.
Moore Jr., Barrington. 1967. *Social origins of dictatorship and democracy*. Boston: Beacon Press.
Moreland, William Harrison. 1923. *From Akbar to Aurangzeb: A study in Indian economic history*. Delhi.
Nicola, Gennaioli, and Hans-Joachim Voth. 2015. State capacity and military conflict. *Review of Economic Studies* 82: 1409–1448.
North, Douglass C. 1982. *Structure and change in economic history*. New York: Norton.
North, Douglass C., and Barry R. Weingast. 1989. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* 49(4): 803–832.
O'Brien, Patrick K. 1988. The political economy of British taxation, 1660–1815. *Economic History Review* 41(1): 1–32.
Oak, Mandar, and Anand V. Swamy. 2012. Myopia or strategic behavior? Indian regimes and the East India Company in late eighteenth century India. *Explorations in Economic History* 49(3): 352–366.
Olson, Mancur. 2000. *Power and prosperity: Outgrowing communist and capitalist dictatorships*. New York: Basic Books.
Perlin, Frank. 1985. State formation reconsidered. *Modern Asian Studies* 19(3): 415–480.
Roy, Tirthanker. 1996. The role of the state in initiating development: A study of interwar South and Southeast Asia. *Indian Economic and Social History Review* 33(4): 373–401.
Roy, Tirthankar. 2012. *Economic history of India, 1857–1947*. Delhi: Oxford University Press.
Roy, Tirthankar. 2013. Rethinking the origins of British India: State formation and military-fiscal undertakings in an eighteenth century world region. *Modern Asian Studies* 47(4): 1125–1156.
Sachs, Jeffrey D. 2012. Government, geography, and growth. The true drivers of economic development. *Foreign Affairs*.https://www.foreignaffairs.com/reviews/review-essay/government-geography-and-growth.
Skinner, G. William. 1977. Introduction: Urban sevelopment in Imperial China. In *The city in late Imperial China*, ed. G. William Skinner, 1–32. Stanford: Stanford University Press. p 19.
Southall, Aidan. 1988. The segmentary state in Africa and Asia. *Comparative Studies in Society and History* 30(1): 52–82.
Stein, Burton. 1985. State formation and economy reconsidered. *Modern Asian Studies* 19(03): 387–413.
Studer, Roman. 2015. *The great divergence reconsidered. Europe, India, and the rise to global economic power*. Cambridge: Cambridge University Press.
Thorner, Daniel. 1966. Marx on India and the Asiatic mode of production. *Contributions to Indian Sociology* 9:33–66.
Tilly, Charles. 1989. Cities and States in Europe, 1000–1800. *Theory and Society* 18(5): 563–584.
Voigtländer, Nico, and Hans Joachim Voth. 2013. The three horsemen of riches: Plague, war, and urbanization in early modern Europe. *The Review of Economic Studies* 80(2): 774–811.
Vries, Peer. 2013. *Escaping poverty: The origins of modern economic growth*. Göttingen: Vandenhoeck & Ruprecht.
Wang, Yeh-chien. 1973. *Land taxation in Imperial China, 1750–1911*, 80. Cambridge, MA: Harvard University Press. 128, 133.

Wink, André. 1986. *Land and sovereignty in India. Agrarian society and politics under the eighteenth century Maratha Svarajya*. Cambridge: Cambridge University Press.
Wolf, Eric R. 1982. *Europe and the people without history*, 86. Berkeley: University of California Press.
Wolfe, Martin. 1972. *The fiscal system of renaissance France*. New Haven: Yale University Press.

Part II
Political Economy of Democracies and Dictatorships

Decentralization, Fiscal Structure, and Local State Capacity in Late-Imperial Russia

Steven Nafziger

Abstract Investments in the fiscal, legal, and infrastructural "capacity" of the state have come to be seen as key determinants of economic development. Central authorities may make these investments, but local public sector institutions also play a role in building state capacity. This chapter examines the interaction between central and local capacity in the context of Tsarist Russia after the end of serfdom. We describe the structure of local government and, drawing on a variety of new sources, provide preliminary evidence on the extent of capacity building by various public sector actors. Our findings are suggestive of a particularly rich interaction between central authorities and decentralized institutions at the local level when it comes to providing public goods and services. We argue that interpretations of early modern and modern state building are remiss if they focus exclusively on the central government without considering the importance of locally determined efforts.

Keywords Mark Harrison • Imperial Russia • Tsar/Tsarist • Mark Dincecco • Paul Gregory • Ministry of Internal Affairs • Ministry of Education • Ministry of Finance • Ministry of the Interior • Peter Lindert • Steven Nafziger • Zemstvo • Mirskie sbory • Mirskie kapitaly • A.P. Fedorov • Nikolai Gogol • Daron Acemoglu

Mark Harrison's research has shaped our understanding of how military strategic concerns, wartime spending, and coercion all influenced the state's role in the economy in the twentieth century, particularly in the Soviet Union (e.g., Gatrell and Harrison 1993; Harrison 1998). Harrison's work in this area touches on a broader scholarship, which argues that perceptions of external and internal threats tend to drive the consolidation of hierarchical control and the development of the capacity

This paper incorporates part of an earlier working paper, "Local Politics and Public Goods in Late-Tsarist Russia." Phonkrit Tanavisarut, Gulya Radjapova, and Ivan Badiniski provided wonderful research assistance.

S. Nafziger (✉)
Department of Economics, Williams College, Williamstown, MA, USA
e-mail: snafzige@williams.edu

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*,
Studies in Economic History, DOI 10.1007/978-981-10-1605-9_3

of central state authorities to collect taxes and enact policies down to the lowest levels of a society. Social scientists have come to consider this development of "state capacity" to be a key feature of political and economic modernization (e.g., Besley and Persson 2010).

Much of the literature on the rise of the early modern state focuses on the revenues and institutions under the control of newly empowered national parliaments or sovereigns (Brewer 1989; Dincecco 2011; Gennaioli and Voth 2015; Hoffman and Rosenthal 1997). Similarly, students of the rise of social welfare states in the later nineteenth and early twentieth centuries tend to emphasize political changes in central institutions and the consequences for revenue and spending policies (e.g., Lindert 2004). But to what extent was development of state capacity over the nineteenth and twentieth centuries only a story of *central* authorities imposing new, additional, or uniform taxes and administrative control on sub-national units? In the case of the Soviet Union, Harrison's pioneering research has described the hierarchical principal-agent relationships that comprised the political and economic systems under Stalin's ultimate control (e.g., Gregory and Harrison 2005). Similarly, Imperial Russia has often been characterized as a highly absolutist and centralized state, subject to the ultimate authority of the Tsar. However, in practice much of the actual governance of the vast empire was local. While state policies and institutions emanating from the center dictated the parameters under which local governments functioned, this chapter argues that a large and particularly influential part of Tsarist "state capacity" was produced and controlled at the sub-provincial level.[1]

Prior to the emancipation of the serfs in 1861, local authority in the Russian empire was firmly in the hands of the landed elite, with only occasional interference by ministerial or military agents acting under central directives. With respect to the post-1861 period, scholars emphasize the failure of the absolutist state to fully centralize control, build robust legal institutions, or develop the capacity to (efficiently) collect revenues and provide economically beneficial public goods and services – i.e., to do precisely what other modernizing states were beginning to accomplish. To a certain extent, this view is true – although the central government was actively involved in supporting infrastructure development, engaging in trade policy, and moving Russia onto the gold standard, below we describe how it collected taxes in a relatively inefficient manner and provided relatively few non-military goods and services, especially ones that impacted the lives of the 85 % of the population who were peasants. However, such interpretations tend to ignore the "capacity-building" that characterized a number of local (sub-provincial) institutions of governance installed in the wake of serf emancipation. As we argue below, these varied forms of local self-government constituted a relatively rigorous form of decentralized governance that at least partially substituted for weaknesses of the central state and the demise of the landed elite.

[1] To some extent, this view is consistent with the framework developed by Acemoğlu et al. (2014) in considering the interaction between local and central state capacity building in modern Colombia.

In recent decades, attention has increasingly focused on decentralization as a mechanism to bypass inefficiencies and corrupt practices of central governments in low-income countries, although one that might lead to increased capture of resources by local elites (Bardhan and Mookerjee 2006). In some cases, decentralization might take the form of intergovernmental grants and the installation of direct central administrative authority at the local level. In other settings, decentralization involves the devolution of fiscal authority and governing powers to local actors. The latter case more closely resembled the Russian context, where peasant communal authorities, all-class representative bodies known as *zemstva*, municipal authorities, and other local institutions received certain legal rights and specific fiscal powers from the Imperial authorities that allowed them to fund a variety of local public goods and services. The resulting local "public sector" of Russia over the period 1860–1917 was a complicated network of ministerial offices, quasi-independent "overseers" of the peasant population, and various rural and urban institutions of self-governance. This chapter focuses solely on Tsarist Russia, but an underlying theme is that the activities and effects of *local* fiscal and political institutions in other societies over the long nineteenth century deserve greater attention from economic historians.[2]

Standard accounts of Russian economic history over the period concentrate on Imperial policies emanating from St. Petersburg.[3] Ministerial decrees, large-scale reforms, and events in the capital cities certainly played a role in Russia's development process, but their impact on the ground, and on the microeconomic decisions of individual actors, was mediated by the institutional structure of local government. This chapter sets out to accomplish two modest tasks. The first is to provide a very basic accounting of decentralized public sector "capacity" in Tsarist Russia in relation to the Imperial fiscal system and central political authorities. Second, we briefly interpret the resulting structure of local and central state capacity in a broadly comparative light. This second task hints at an underlying methodological concern of this chapter – although we only study the Russian case, this example can offer insights into state capacity, decentralization, and fiscal systems in developing countries and in the rise of the modern nation state. Given the speculative nature of the evidence and claims put forth here, it is very clear that considerable work remains to establish the role of local state capacity in the process of economic development in Russia and elsewhere over the long nineteenth century.

[2]Important recent works on European and global fiscal development prior to World War I, such as the chapters in Cardoso and Lains, eds. (2013) and Yun-Casalilla and O'Brien, eds. (2012), barely touch upon sub-national components of state capacity.

[3]The little work that does exist on Russian institutions of local government tends to rely on commentary and legal decrees emanating from the center. For example, see Lapteva (1998) and Starr (1972).

State Capacity in Imperial Russia: An Overview

Recent scholarship in political economy has fixed on the development of the *means* of enacting pro-growth policies as a critical step in the process of economic development. In practice, definitions of this "state capacity" tend to focus on central authorities. Studies such as Dincecco and Katz (2016) document state capacity as per capita central government revenues (extractive capacity) or non-military spending (productive capacity).[4] Besley and Persson (2010) consider various measures of legal capacity, such as protection of property rights and financial development, and of fiscal capacity, including the amount of non-tariff revenues and various types of taxes imposed by central governments. Here, we primarily focus on fiscal dimensions of state capacity in the Russian context, with brief sidebars on legal and "infrastructural" measures.

In various works on the Tsarist transition to modern economic growth, Paul Gregory de-emphasized the role of the state's fiscal practices in overcoming various limitations of Russia's economic backwardness. Gregory's (1982—also see Table 1) NNP estimates suggest that that the share of total government spending (excluding capital investments) was relatively large in fluctuating between 8 % (644 million rubles) in 1885 and 11 % (2.2 billion rubles) in 1913, although he strongly argues that total spending may not have translated into a large role for the central government in the development process. Gregory noted that the share of military spending in total government expenditures was quite large over the period – 45.6 % in 1885 to 44.9 % in 1913 – and that this, along with other aspects of the state's involvement in the economy, did very little to improve the allocation of resources, generate additional investment, provide public goods, or spark innovation.

Figure 1 summarizes Russian central government spending over the period 1804–1913. These data are based on amounts spent through Tsarist ministries, with some extraordinary expenditures included.[5] As such, any categorizations miss relevant expenditures that occur through largely unrelated ministries (e.g., educational expenditures taking place through the Holy Synod). Overall, while changing values of the ruble are obvious concerns, spending picks up after 1870. The share of non-military expenditures slowly increased, although our measure likely understates military spending (defined as the sum of spending by the army and naval ministries). Expenditures on internal governance (mostly through the Ministry

[4]Gennaioli and Voth (2015) utilize total central government revenues as their main measure of state capacity.

[5]Gregory's (1982) measures include not only the expenditures of the central government, but also those of municipal authorities, the *zemstvo*, and local institutions of peasant self-government. That Gregory works with net government spending (net of transfers and intermediate purchases) likely explains the difference between his numbers and ours, both for central government expenditures and those of local institutions. On the structures of government in the Imperial period, see Hartley (2006) and Shakibi (2006).

Table 1 Russian state capacity in comparative perspective

	1885	1900	1913
Population (Empire)	109 million	133 million	171 million
Net national product (Gregory)	8 billion	13.4 billion	20.2 billion
Net total government spending (Gregory)	644 million	1.12 billion	2.22 billion
Net local government spending (Gregory)	145 million	294 million	643 million
Total government spending/NNP (Gregory)	8.05 %	8.36 %	11.00 %
Local government spending/NNP (Gregory)	1.82 %	2.19 %	3.18 %
Current paper – totals in current rubles			
Central government revenues	833.9 million	1.74 billion	3.43 billion
As share of NNP	10.40 %	13.00 %	17.00 %
Direct tax receipts	105.1 million	62.1 million	272.5 million
As share of NNP	1.31 %	0.46 %	1.35 %
Indirect taxes (including alcohol)	290.4 million	641.9 million	1.25 billion
All alcohol revenues	231.2 million	434.7 million	953.0 million
Tariff revenues	95.0 million	204.0 million	352.9 million
Loans and bond issues	71.6 million	32.6 million	13.8 million
Comparative evidence – total central government revenues/GDP			
France	12.19 %	11.65 %	10.29 %
Germany	1.97 %	2.73 %	3.19 %
Italy	13.08 %	12.85 %	11.82 %
Spain	8.29 %	8.87 %	9.35 %
Sweden	6.52 %	6.91 %	6.80 %
United Kingdom	7.32 %	7.82 %	8.43 %
Comparative evidence – direct taxes/GDP			
France	1.71 %	1.60 %	1.28 %
Germany	0.00 %	0.00 %	0.00 %
Italy	1.78 %	3.72 %	2.43 %
Spain	2.37 %	2.47 %	3.30 %
Sweden	0.81 %	0.41 %	1.06 %
United Kingdom	1.46 %	1.62 %	2.13 %

All values are in nominal (silver) rubles unless otherwise noted. The sources of the Russian data are Gregory (1982), or are defined in the text and under Fig. 1. The comparative evidence comes from the Global Price and Income History project (gpih.ucdavis.edu)

of Internal Affairs) hovered between 5 % and 10 %, while Ministry of Education spending never consisted of more than 6 % of the central budget.

The more commonly used indicators of state capacity lie on the revenue side of the budget. Of course, total revenues including debt financing closely paralleled the long-run rise in expenditures. However, the composition of these revenues changed in important ways over time. Figure 2 presents such information. Overall, the bulk of central revenues between 1864 and 1914 came from indirect taxes (mean of 25.6 %),

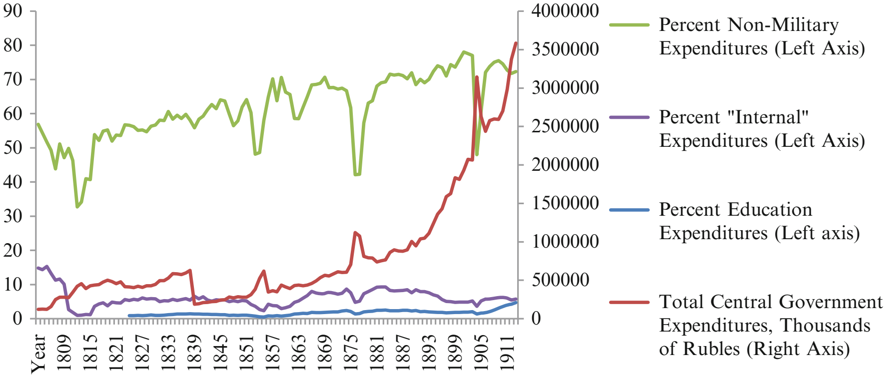


Fig. 1 Central government spending, 1804–1914. Note: The percent of non-military spending excludes naval, army, and extraordinary expenditures explicitly earmarked for military purposes. The percent of "Internal" expenditures includes ordinary spending earmarked for the Ministry of Internal Affairs and the postal system. The percent of expenditures on education refers to the Ministry of "Enlightenment," exclusively, although educational spending did occur through other ministries. Non-military spending includes Education and "Internal" expenditures. The data from 1804 to 1902 come from Rozhdestvenskii, ed. (1902). The remaining years of data are drawn from the relevant Yearbooks (*Ezhegodniki*) of the Ministry of Finance. Except for the years 1905 and 1914, all of these data represent actual expenditures rather than budgeted amounts

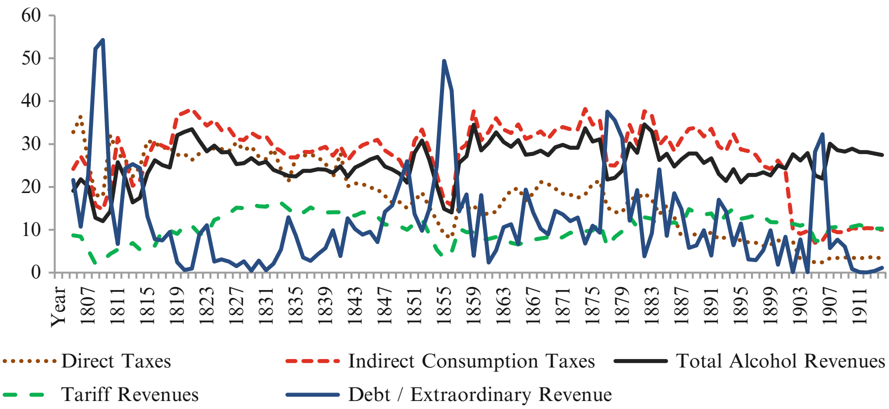


Fig. 2 Shares of central government revenues. Note: All series represent percentages of total central government revenues, which include ordinary, one-time, and extraordinary (overwhelmingly debt financed) revenues. "Direct Taxes" includes the soul tax, land taxes, taxes (patents) on trade rights, and industrial profit taxes. "Indirect Consumption Taxes" include various sales taxes, including excise taxes on alcohol, salt, sugar, tobacco, oil, and other commodities. "Total Alcohol Revenues" include the same alcohol excise tax AND (after 1895) direct production revenues and retail sales. "Debt/Extraordinary" revenues are almost entirely foreign and domestic bank loans and bond issues. The data sources are the same as in Fig. 1

tariffs (10.6 %), and, especially, alcohol taxes and sales (27 %).[6] Direct head (soul), property, and industrial property taxes only comprised 11.2 % of total revenues over this period, with debt payments constituting a volatile but significant share at 11 %.[7] Significantly, most direct taxes were imposed collectively at a relatively high level (typically), with the subsequent allocation to lower units determined largely at the local level (from *zemstvo* on down), who included their own obligations alongside the central government's demands (Kotsonis 2014).

To make comparisons to the broader literature on state capacity, these revenue numbers should be scaled by the size of the Russian economy. Gregory's NNP numbers (and government spending shares) are available only for the period 1885–1913. Therefore, we limit our comparisons in Table 1 to just 3 years: 1885, 1900, and 1913. Over those 3 years, Russian state capacity as measured by central government revenues relative to NNP rose to the highest level among comparable European economies (17 %). As reflected in Fig. 2 as well, this rise was largely dependent on indirect sources – excise taxes, tariffs, and alcohol revenues. Thus, while the Tsarist state did appear to possess considerable revenue-getting capacity, there was little reliance on more structured sources emphasized by Besley and Persson (2010). Indeed, discussions of broader income taxes went nowhere before 1917 (Lindert and Nafziger 2014), although there were some inroads into business and corporate income taxation (Kotsonis 2014). Moreover, the record of Tsarist central government revenues may be misleading if the main center of Russian state capacity resided in the institutions of the local public sector, which we now move on to describe.

Institutions of Peasant Self-Government

The bottom rung in the Tsarist governance structure comprised the institutions of local peasant self-government: the new versions of the commune (the "rural society," or *sel'skoe obshchestvo*), the township administration (*volost'noe pravlenie*), and associated township courts and local police, which all received formal stature in the emancipation reforms of the 1860s (Gaudin 2007). Informal versions of such institutions had existed for hundreds of years, and an enormous historiography

[6] These indirect taxes included excise (*aktsiz*) taxes on other goods (tobacco, sugar, etc.); patents granting the right to sell these same goods; direct state production and sales of alcohol (after 1895); various fines and fees; ticket sales and fees on state railways; and others.

[7] As the state treasury became increasingly reliant on indirect taxation over the period, it appears that the bulk of direct tax revenues were left in the hands of local offices of central authorities (Zakharov et al. 2006). The composition of direct taxes changes over time, with the cessation of the soul tax in 1886 and its replacement by a state land tax. Direct taxes included other forms of property taxes as well.

describes the commune and communal governance across space and time before 1861.[8] Earlier reforms of the state (non-serf) peasantry in the 1830s established a nested township-commune system and prescribed a set of responsibilities, such as the maintenance of grain stores, the collective fulfillment of external taxes, and the execution of so-called "natural" obligations that included supplying recruits and upkeep on local roads. This served as a model for a common structure of peasant self-government established over much of the empire after 1861.

Peace mediators (*mirovye posredniki*) and new district and provincial administrations of peasant affairs (*krest'ianskoe po delam prisutstvie*) under the Ministry of Internal Affairs managed the process of setting up these institutions of self-government. Townships were supposed to include no more than 3000 (male) souls, and so the underlying number of constituent rural societies varied widely across Russia from one to dozens. Decisions made by peasant communes and townships emerged out of meetings of "representatives" of the peasant population that were presided over by elected elders.[9] Communal elders, township elders, township clerks, and other employees of peasant local governments comprised the bottom layer "officialdom" in Tsarist Russia. State and *zemstvo* officials often called upon them to execute policies and to report on local conditions. Critical to these efforts were paid police deputies (*sotki* and *desiatki*) selected from among the peasant population. In 1880, there were approximately 10 such local police per 1000 rural residents across European Russia (Russia Ministerstvo vnutrennykh 1881).[10] These were in addition to locally stationed, MVD-employed constables and their staffs.

The newly formalized peasant communes and townships could also assign local "taxes" (*mirskie sbory*) on members to hire staff and support a variety of other public goods and services. Such collections were often made alongside external obligations to the *zemstvo* and central government, with the rural society collectively liable for submitting payments two times per year. Townships and rural societies also managed grain stores and cash funds, making loans and imposing supplemental

[8]For a summary, see Mironov (1985). Prior to 1861, peasant and urban leaders occasionally assessed community members to provide some services, such as paying a literate villager to teach in an informal school. However, historians of serfdom have found little evidence of significant welfare or public good provision by serf communes (Dennison 2011). The Ministry of State Domains, which administered (and collected revenues from) the state peasantry, did establish a grain storage network, founded primary and secondary schools, and organized rural health networks after 1830. These were rather limited efforts, but they did provide examples followed by other ministries and, later, by the *zemstva*. On public good provision among the state peasants, see Ivanov, (1945). For discussions of urban government and public service provision prior to 1861, see Brower (1990).

[9]By "representatives" in these peasant institutions, we are referring to household heads in the communal *skhody*, or assembly, and the community members and rural society elders sent to attend township-level *skhody*.

[10]Other paid employees of rural societies and townships included tax collectors, guardsmen (over grain stores and churches), and agricultural workers such as shepherds for community flocks.

Table 2 Township and communal expenditures in 1881

Budget categories	46 provinces of European Russia Rubles	% of total
Salaries for *Volost*’ elders	1,865,441	12.37
Salaries for clerks/other office employees	3,300,393	21.88
Maintenance of *Volost*’ structures	1,526,234	10.12
Volost’ office expenses	236,484	1.57
Maintenance of local churches	191,792	1.27
Spending on education	900,027	5.97
Spending on health care	371,619	2.46
Spending on public assistance	39,588	0.26
Spending on fire prevention	128,015	0.85
Spending on food relief	127,831	0.85
Spending on road maintenance	141,648	0.94
Spending on other transportation	76,295	0.51
Spending on troop quartering	316,138	2.10
Spending on horses	3,740,674	24.80
Salaries of guards/watchmen	63,346	0.42
Spending on elections for juries and *Zemstv*a	42,302	0.28
Spending on arrests and detainment	26,629	0.18
Other spending	1,991,917	13.20
Total *Volost*’ spending	15,086,373	
Salaries for rural society elders	2,157,926	12.37
Salaries for other rural society employees	5,618,529	32.21
Other spending	9,667,012	55.42
Total rural society (commune) spending	17,443,467	

Source is Russia, Tsentral’nyi (1883). Data exclude the Baltic provinces and the Don Cossack Land

collections when necessary.[11] In the absence of debt markets or credit supplies, revenue and expenditure totals of these bodies were essentially equal, year-by-year. Focusing on the latter, Table 2 reports spending by peasant institutions in 1881 (1905 data show a similar pattern). Salaries for clerks and elders and the maintenance of structures, offices, and horses took up much of township spending in 1881, and a similar distribution – although focused on somewhat different public goods, especially schools – is evident for rural societies (not fully reported here).[12] Especially telling is the wide variety of activities that these local institutions could

[11]Documentation of the grain storage system is widely available among the archival holdings of *zemstvo*, peasant institutions, and local Offices of Peasant Affairs. Such records include account books – see GANO, 20.90.46.

[12]Township elders were paid roughly 200 rubles per year in both 1881 and 1893, while rural society elders received approximately 30–40 rubles (1893 data not reported here). Township clerks – much more likely to be literate – were generally paid more than the elders.

and did engage in. Moreover, many in-kind services such as local fire protection and the organization of customary law courts are not indicated in these budget totals.

Figure 3 explores the geography of spending by these institutions in 1881. While the levels were not terribly high (0.26 rubles per [rural] capita by townships; 0.31 by rural societies), township spending was relatively greater in the non-*zemstvo* (see below) region, while rural society spending was higher in the provinces to the north and east of Moscow. The former likely reflected additional obligations in the absence of the *zemstvo* in the western provinces. Overall, these maps indicate considerable variation in the level of activity undertaken by these local institutions, the implications of which have yet to be explored.

The *mirskie sbory*, allocated along with shares of external obligations among households by the communal and township assemblies (often by the number of "obligated souls"), constituted the main but not only source of funding for these institutions. Table 3 provides some indication of this, as "other" sources constituted approximately 25 % of total revenues in Morozov township (Moscow province). Archival evidence suggests that many peasant institutions built up cash "communal capital" (*mirskie kapitaly* – often from the sale or rental of communal assets) on deposit at local financial institutions or with the township, and that these assets generated returns. Rural societies could rent out portions of the land received in the emancipation settlements, and they could charge for the right to run drinking and eating establishments. Township authorities collected fees for issuing work and travel passports to the local population.[13] Total fiscal demands (excluding any in-kind or labor obligations) placed on members by rural societies and townships tended to be less than external state burdens, but this locally generated "capacity" did support schools, local policing, and other public services to an underappreciated degree.

The *Zemstvo*

At noon on October 23, 1883, district *zemstvo* executive committee chairman A.P. Fedorov and 31 assemblymen filed into the district courtroom in the town of Ardatov in Simbirsk Province to decide on a series of budget issues and to vote over a new executive for the next 3-year term.[14] With the issuance of the *Statutes on Provincial and District Zemstvo Institutions* in 1864, Tsarist reformers explicitly viewed such assemblies – comprising representatives of urban, rural, and peasant communal property owners – as constituting a mechanism for *all-class* self-

[13]This is evident in numerous archival records that provide rough financial accounts of specific rural societies and township. For examples, see TsIAM, 199.2.362; TsGIA SPb, 190.5.286; and GANO, 20.90.113b.

[14]This description of Ardatovskii district *zemstvo* activities for 1883 is taken from minutes published in *Zhurnaly* (1884, pp. 106–223).

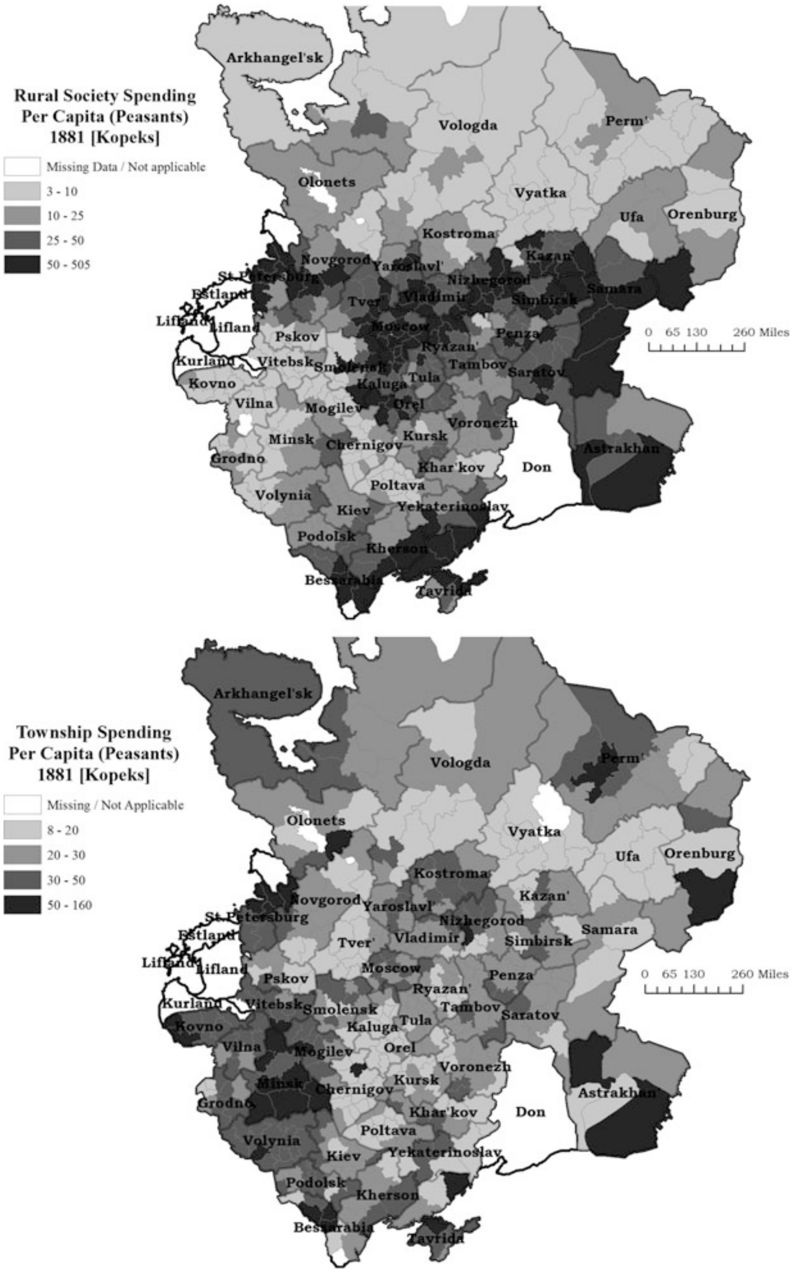


Fig. 3 Rural society and township spending per capita, 1881. Note: The source – including the underlying population totals – is Russia, Tsentral'nyi (1883)

Table 3 Rural societies in Morozovskaia Township, Dmitrov Fistrict, Moscow Province – 1892

Rural society	Total population	Obligated "Souls"	*Mirskie Sbory*	Other revenue	Elder's salary
Morozovskoe	1409	579	1272.35	1280	150
Rakhmanovskoe	650	269	813	143	150
Fedorovskoe	1103	492	1257.75	238	175
Putilovskoe	1116	512	1164.5	422	177
Muromtsevskoe	998	456	1090.4	150	162
Novlenskoe	588	233	823	0	115
Gevlevskoe	254	45	300.75	0	70
Bulakovskoe	265	71	341.9	0	60
Spasskoe	307	88	550.4	0	70
Mashinskoe	226	54	236.6	0	70
Shelkovskoe	368	97	689.1	0	100
Vysokovskoe	981	361	1265	324	160
Klement'evskoe	721	248	630.3	906	140
Deulinskoe	726	267	882.75	100	150
Total	9712	3772	1,1317.8	3563	1749

Source: TsIAM 199.2.362. Two of these rural societies mention support for schools

government in the 34 (eventually 40) provinces of European Russia where they were established.[15] Under the law of 1864 (comprehensively revised in 1890), different types of property owners – private, urban, and peasant communal – were assigned specific numbers of seats in each district *zemstvo* assembly.[16] By providing for explicit representation from the peasantry, the *zemstvo* was at least partially intended as a way to deal with rural needs that might have exceeded the capabilities of the local institutions of peasant self-government.[17]

The initial act required the *zemstva* (pl.) to finance other local institutions (such as district courts above the township ones), to manage military provisions and

[15]See *PSZ* (Series II, Vol. 39, 1864, No. 40457). In this way, the *zemstva* have been viewed as a response to the "problem of provincial under-institutionalization" in Tsarist Russia (Robbins 1987, p. 16).

[16]Discussions over the original *zemstvo* law and the 1890 reform cited population and the distribution of property as the key variables behind the setting of seat shares. The reformers explicitly acknowledged the intent to favor the local landed elite as the most educated and experienced people in the provinces. In addition, the Minister of the Interior, P. A. Valuev, in his proposal for the 1864 law, outlined district norms (*tsenzi*) of communal or private land that were meant to correspond to curia seats in the two curia (the urban curia seats were to be based primarily on population).

[17]These assemblymen (*glasnye*) elected the district executive boards and representatives to provincial assemblies (which then elected a provincial executive committee). Conservative reforms of the 1890s reduced the assembly shares of the peasant and urban curiae. However, the newly emancipated peasantry still retained seats in the *zemstvo* assemblies and the possibility of election to executive positions. For additional detail, see Nafziger (2011).

Table 4 Total *zemstvo* revenues and expenditures, select years

	Revenues			Expenditures			
	Property taxes	% of total	Total	Education	Health care	% of total	Total
1871	15.6	72.7	21.5	1.6	2.1	18.1	20.7
1880	26.8	73.7	36.3	5.0	6.4	32.6	35.1
1886	28.1	67.8	41.5	6.7	9.2	36.7	43.4
1896	42.3	70.4	60.1	9.9	18.3	46.3	60.9
1903	64.6	64.9	99.5	19.1	30.2	49.6	99.5
1906	83.4	67.1	124.2	25.3	35.9	49.3	124.2
1913	155.4	62.3	249.4	87.7	70.2	63.3	249.5

Numbers refer to the sum of district and provincial revenues and expenditures for just the 34 provinces with *zemstva* in 1903 in millions of current rubles. The spending totals for 1871 and 1880 do not include Samara province. Property tax income in 1871 and 1880 is defined slightly more broadly than the years that follow. Data for 1871 and 1880 are budgeted rather than actual totals. Finally, the difference in total spending and income for 1913 reflects extra expenditures on items budgeted in 1912 (Source: Nafziger 2011)

grain stores, to maintain roads and communication networks, and to aid in the collection of taxes for the central government. In addition to these "obligatory" responsibilities, the founding statutes called on the *zemstva* to undertake programs to support "the local economic and welfare needs of each province."[18] Over the following half century, this mandate led to substantial *zemstvo* involvement in the expansion of rural education and health care, in the support of local artisans and craftsmen, in encouraging credit and cooperative organizations, and in providing veterinary and agronomic services to farmers.

The scale of the *zemstvo*'s growing role in public health and education is suggested by the increasing expenditures noted in Table 4 and by the decomposition of aggregate *zemstvo* expenditures in 1883 in Table 5.[19] Even considered in isolation, the expenditures depicted in Table 5 suggest that in the provinces where provincial and district-level *zemstvo* existed, they were key components of local government. We depict the geography of *zemstvo* expenditures per capita in 1906 (the picture for revenues per capita is similar) in Fig. 4. There was considerable variation in *zemstvo* spending across districts, expenditures were slightly higher in northern and northeastern provinces, and the level of spending per capita was roughly of the same magnitude as spending by the local institutions of peasant self-government.

After meeting for 3 days, the Ardatov district *zemstvo* assemblymen were ready to hear final reports on issues ranging from the ongoing construction of a village school to the *zemstvo*'s activities in road maintenance over the past year. Two

[18] *PSZ*, Series II, Vol. 39, 1864, No. 40457, Clause 1.

[19] As reported in *Zhurnaly*, the Ardatov budget for 1884 included 81,481.64 rubles in expenditures, with 31,756.96 for health care (including the salaries and expenses for four doctors and three small hospitals) and 12,139.30 for education (including 5160 rubles in salary for 35 teachers).

Table 5 Total provincial and district *zemstvo* expenditures in 1883

Budget items	Expenditures (rubles)	% of total
Supporting other local administrative institutions	1,587,331	4.23
Supporting local judicial institutions	4,390,695	11.71
Supporting local Bureaus of Peasant Affairs	1,356,107	3.62
Expenditures on *Zemstvo* administration	4,323,580	11.53
Medical expenditures	8,497,249	22.67
Education expenditures	6,098,186	16.27
Roadwork/infrastructure	2,665,140	7.11
Military provisions and housing	4,207,424	11.22
Public assistance	2,312,994	6.17
Combatting pests	40,830	0.11
Food relief/supply	132,503	0.35
Debt payments	410,888	1.10
Indirect expenditures (transfers to other *Zemstvo*)	437,013	1.17
Other expenditures	1,023,352	2.73
Total	37,483,292	

Source of the data is Russia, Tsentral'nyi (1886)

pieces of business were especially important. First, the assembly heard a report on the planned budget for 1884. The proposed budget projected 81,521.15 rubles in revenues for the 1884 calendar year, 65,140.68 (79.9 %) of which derived from a tax of 12.08 % on the estimated (yearly) income generated by land and other "immovable" property in the district.[20] Across European Russia, property owners complained about such tax assessments, and the *zemstva* were forced to rely on local police, urban officials, and peasant governments to aid in the collection of *zemstvo*, state, and local obligations.[21] According to the law, *zemstvo* obligations received a lower priority than did direct state taxes (referred to in the *zemstvo* documentation as *kazennye sbory*, which included soul, land, and other direct taxes) in the final allocations.[22] Beyond property taxes, *zemstvo* also held the right to collect payments for trade and commercial rights, for the usage of *zemstvo* property, and for various services that they provided (including medical care). However, as Table 4 indicates,

[20]As a point of comparison, by the early 1880s, the *volosti* and *sel'skie obshchestva* of Ardatov district were spending roughly 120,000 rubles in total (Russia, Tsentral'nyi 1886).

[21]For example, a substantial part of the 1890 business of the Semenov district *zemstvo* executive board (*uprava*) was taken up with efforts to deal with tax complaints and arrears (GANO, 51.251.292).

[22]For additional details, see Nafziger (2011) and Russia, Tsentral'nyi (1896). In 1896, Ardatov district *zemstvo* expected to collect 70,421 rubles from property taxes but only received 55,396. Such shortfalls resulted in total accumulated arrears on property taxes of 83,300 rubles by the end of 1897. To help finance this deficit, by January of 1898 the *zemstvo* had borrowed 42,469 rubles against its capital reserves and 15483 (at 4.5 % interest) from the provincial *zemstvo*. Despite this persistent gap, budgeted spending rose from 103,000 rubles in 1896 to 137,000 in 1903. See ibid.; and Russia, Statisticheskoe (1906 volume).

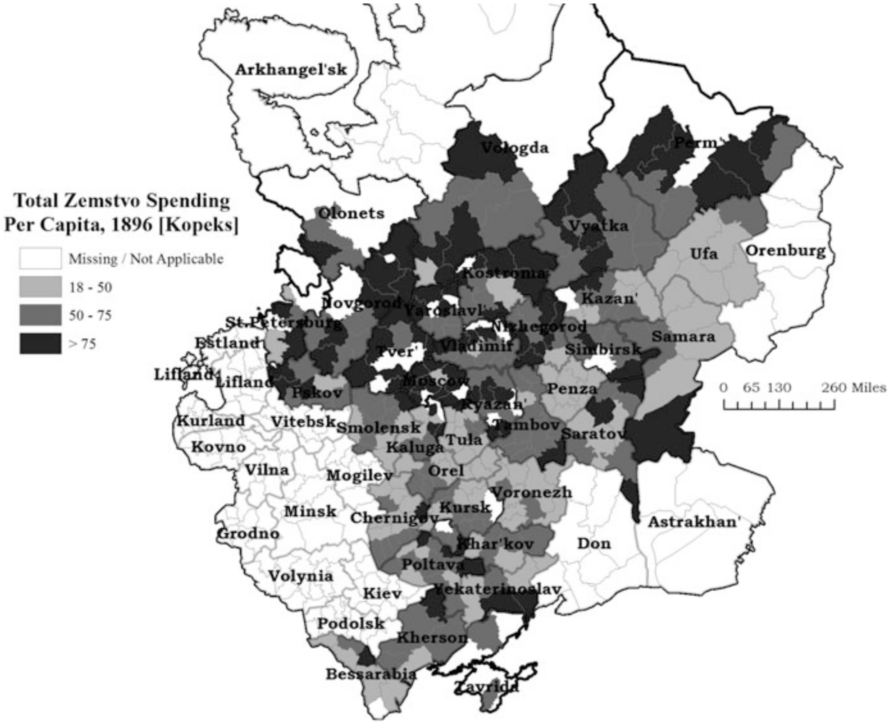


Fig. 4 Total *zemstvo* expenditures per capita, 1906. Note: These data sum over provincial and district *zemstvo* expenditures. For the sources, see Nafziger (2011). The map of revenues per capita looks almost identical. The underlying population totals only refer to the rural population

property taxes contributed roughly 70 % of total revenues throughout the period across all *zemstvo*. Only towards the end of the period did grants from the central government constitute any substantial part of *zemstvo* revenues.[23]

After the budget was approved by majority vote, the Ardatov assembly came to their second piece of important business: voting over executive positions for the 1883–1886 electoral period. These included a new executive board chairman, new executive board members, two *zemstvo* representatives to the district school council, and nine representatives to the Simbirsk provincial *zemstvo* assembly. After a long series of nominations and votes, Filipp Mikhailovich Mikhailov, a peasant, was elected as the executive board chairman, two other peasants – Mikhail Timofeevich Diagilev and Petr Vasil'evich Turgenev – were voted in as the other members of the executive committee, and peasants filled all nine district representative positions in the provincial assembly. As a result, executive power in the Ardatov *zemstvo* lay entirely in the hands of the peasantry, even though by statute the peasant electoral

[23]The majority of these grants were matching funds tied to school building.

curia only comprised 23 of the 52 seats in the district assembly. This contrasted with most other *zemstva*, where election outcomes in 1883 tended to reflect the weighting of assembly seats toward the curia of rural private property owners (and the nobility).

Ardatov's *zemstvo* had a relatively large share of peasants elected to positions of authority, and the district seems to have spent a comparatively large share of its budget on health care and education. At the same time, the mean *zemstvo* tax rate on the 51 % of district land that was communally owned was 24.3 kopeks per *desiatina* (2.7 acres) in 1885, versus 22 kopeks for the 32 % of land in private hands, despite grain yields being higher on privately owned land.[24] Comparing 1883 data on electoral outcomes to the 1885 tax data and other information, we can more formally evaluate whether there was a relationship between peasant political power in the district *zemstvo* and the gap in tax rates between the two types of property (see Nafziger 2011). The econometric results indicate that the districts where peasant assemblymen achieved more political power showed lowered property tax gaps. However, we cannot entirely discount the possibility that those districts where peasants received a greater political voice had some unobservable characteristics that were correlated with differences in land quality between the two types of property.[25] What does seem to be clear is that the *zemstvo* assemblies and executive committees were quite active in fostering fiscal capacity that reflected and catered to local needs.[26] Among other effects, these investments translated into real improvements in school provision in the countryside in the absence of substantial involvement by the central authorities until very late in the period.[27]

For the non-*zemstvo* provinces of European Russia, some of the *zemstvo* functions were undertaken by the local institutions of peasant self-government, often in concert with the local offices of peasant affairs and other representatives of the central government. However, officials under the direct authority of the provincial governors made the majority of local funding and revenue decisions in these districts.[28] In 1905, the Ministry of Finance recorded 24.5 million rubles in *zemskie* revenues in the non-*zemstvo* provinces, with over half (14.5 million) of the

[24]For the tax rates and property allocation, see Skanlon, ed. (1888). For information on the *zemstvo* electoral outcomes of 1883, see Syrnev, A. ed. (1888).

[25]Nafziger (2011) goes on to investigate the relationship between peasant representation and expenditures in more depth by relying on a change in the composition of *zemstvo* assemblies after a reform of 1890. That paper finds evidence consistent with peasant influence in these institutions.

[26]On the breakdown between provincial and district *zemstvo* activities, see Veselovskii (1909, vol. 1).

[27]The supply of primary schooling grew faster in *zemstvo* provinces than non-*zemstvo* ones, even after controlling for a variety of other possible explanations. *Zemstvo* efforts in health care, in promoting rural industry and crafts, in providing veterinary and agronomic services, and in managing large-scale fire insurance systems were critical components of a developing rural service sector.

[28]The specific official in charge differed depending on the region (Lapteva 1998). For details on revenue sources and expenditures undertaken by these and other local officials in non-*zemstvo* areas, see various yearbooks of the Ministry of Finance.

receipts from local property taxes and roughly two million in direct grants from the central government.[29] Overall, the fiscal and governance structure of the non-*zemstvo* provinces deserves further attention.

Where they existed, the *zemstvo* amounted to a form of "decentralization from scratch," installed as a mechanism to build local state capacity and to translate the resulting resources into local public services. The peasant institutions of self-government played an important role as the locus for electing peasant representatives to the *zemstva*, while also functioning as an independent base of local state capacity on their own. Thus, while the central authorities were not completely absent, it is clear that a significant share of funding for local public goods in rural European Russia was locally collected and controlled.

Local Corporate and Municipal Governance

Social classes other than the peasantry also maintained their own local institutions of self-government. These included merchant guilds, district assemblies of the nobility, and townsmen (*meshchane*) associations. While their roles ebbed over the late Imperial period, these bodies continued to be called upon to administer aspects of local governance and contribute (often charitably) to schools and other public goods and services.[30] In this sense, these institutions did contribute to the building of formal state capacity at the local level. Unfortunately, budgetary information from these bodies is largely unavailable, as accounting was typically informal, and records were rarely kept. As such, the relative importance of these institutions in the system of Tsarist local governance remains largely unknown.

While the Russian Empire remained overwhelmingly rural, urban governments were an increasingly important part of the state structure over the last decades of the Tsarist era.[31] Assemblies and mayoral forms of urban governments were allowed in formally chartered settlements, and after reforms in the 1860s, such town authorities held the right to assess taxes on urban property, to collect certain fees and patents, and even to issue debt. Funds went to support local administration, schools, public health, and welfare provision of various types. By 1912, urban governments were

[29]Thus, the underlying funding sources in these non-*zemstvo* provinces were not so different in practice from what the *zemstvo* had available. It is not clear precisely where these two million rubles show up in the central government budget – the expenditures of the Ministry of Internal Affairs in 1905 amounted to 114.4 million rubles, with roughly 76 million dedicated to "local administration" and approximately one million separately appropriated for *zemstvo*, municipal, and other local institutions (Russia, Ministerstvo finansov 1907).

[30]For example, land captains (see below) were formally drawn from the district assemblies of the nobility, and townsmen associations played a role in municipal electoral systems. These bodies assessed obligations on their members to fund various initiatives (Hamburg 1984; Rieber 1991; Wirtshafter 1997).

[31]On urban government in Imperial Russia, see Brower (1990) and Koshman (2008).

spending approximately 275 million rubles, which is comparable to the aggregate expenditures of the *zemstvo*.[32] As Table 6 indicates, the majority of the associated revenues came from property taxes and various fees charged to urban enterprises and users of public services. The implication of such evidence – evident for earlier years as well – is that considerable local state (fiscal) capacity emerged in cities to match what the *zemstvo* and other rural institutions were doing.

Central State Capacity at the Local Level[33]

The central authorities were not completely absent from local governance. Although scholars (e.g., Starr 1972) do assert that the central government employed relatively few officials outside of the provincial capitals, the truth was far from Nikolai Gogol's government inspector, whose visit to the provincial town was so unexpected that it caused pandemonium among local residents. In practice, a variety of state officials impacted the workings of the institutions of local government on a day-to-day basis. The politics and policymaking of the *zemstvo* and the peasant institutions of self-government were embedded in a complex structure of centrally controlled judicial and supervisory authorities at the district and provincial levels. These included the peace mediators that managed the emancipation process (Easley 2008); governors, police, and other administrative bodies under the Ministry of the Interior; local treasury offices; district courts and justices of the peace; and the land captains (*zemskie nachal'niki* – after 1889). The funding for such entities came from a combination of retained local sources and transfers from the central ministries, although it has proven impossible to credibly identify the local components of Ministerial budgets.

The most prominent "local" officials were the provincial governors and their staffs.[34] These appointees of the central government possessed significant executive power (including the authority to call out troops) but no direct independent control over the level of taxes or legal capacity in their provinces. Moreover, they controlled no specific budget to fund local public goods and services. However, they did possess final approval over *zemstvo* budgets and policies, a right that was

[32]Urban spending rose from 38 million rubles in 1880 to 56 million in 1890, before increasing even more sharply over the next two decades. In 1912, revenues of approximately 13 rubles per capita – much higher than what other government institutions collected from their constituents – supported this spending. See Russia, Ministerstvo finansov (various).

[33]I touch on only a small number of issues here. Further work is necessary to fully document the interrelationship between central and local state capacity in this period. For some general insights in English, see Waldron (2007, Part 2) and Yaney (1973, esp. Chap. 9).

[34]Robbins (1987) and others have documented the characteristics and impact of the largely noble-class governors, noting their particular careerist concerns.

Table 6 Urban government revenues and expenditures, 1912

		European Russia (with Poland and Caucuses)	Whole empire
Number of Cities	Incorporated	678	775
	Non-incorporated	130	196
Total Revenues (thousands of rubles)	Incorporated	216,280.1	24,7458.6
	Non-incorporated	24,756.5	28,221.4
Revenues per resident (rubles)	Incorporated	12.92	12.99
	Non-incorporated	8.27	7.04
Main categories of revenues (thousands of rubles)			
Collections from immovable property	All	28,727.5	32,673.6
Collections from trade and industry	All	13,190.3	14,506.5
Fines of different types	All	3055.6	3340.7
From urban property (renting/usage)	All	40,660.5	46,615.8
Fees from urban enterprises (*sooruzhenii*)	All	75,341.8	79,641.4
Fees for services	All	44,743.4	55,987.6
Expenditures, total (thousands)	Incorporated		247,398.4
	Non-incorporated		28,551.4
Percentages of all expenditures			
Debt payments	Incorporated		15.5
	Non-incorporated		12.1
Supporting government enterprises	Incorporated		15.4
	Non-incorporated		4.9
Medical, veterinary, sanitation	Incorporated		12.9
	Non-incorporated		9.2
Education	Incorporated		12.9
	Non-incorporated		6
Welfare provision	Incorporated		7.5
	Non-incorporated		26.1
Upkeep of government property	Incorporated		7.5
	Non-incorporated		17.9
Upkeep of administration/courts	Incorporated		7.4
	Non-incorporated		7.3
Support of urban police	Incorporated		5.7
	Non-incorporated		7.6
Military billeting	Incorporated		4.2
	Non-incorporated		0.6
Social charities	Incorporated		3.5
	Non-incorporated		0.5
Fire measures/department	Incorporated		2.7
	Non-incorporated		1.8
Various expenses	Incorporated		1.6
	Non-incorporated		2.7
Support for other state institutions	Incorporated		1.1
	Non-incorporated		2.1
Tax payments	Incorporated		1.1
	Non-incorporated		0.4
Payments into educational capital	Incorporated		1
	Non-incorporated		0.8

These data come from collected urban budgets for 1912 (Russia, Departament, 1917) strengthened in 1890, and they could employ police power to intervene in other local institutions.[35]

The district and provincial offices of peasant affairs (*krestian'skie po delam prisutstviia*) were one of the most prominent state authorities in rural areas. Formally under the Ministry of Internal Affairs, these offices were established in 1874 to take over some of the functions of the peace mediators (*PSZ*, 2nd series, vol. 49, no. 53678). Staffed by a combination of local residents and appointees from the center, they were responsible for monitoring the activities of the *zemstvo* and the rural societies and townships to make sure statutes were followed, for responding to complaints from peasants about the functioning of those institutions, for communicating dictates from central authorities, and for managing the electoral processes for the *zemstvo* and peasant institutions. Pearson (1989) notes that a key role of these bodies was to monitor the collection of different levels (*mirskie*, *zemstvo*, or central) of assessed property taxes. He also argues that over time these offices saw their roles grow as part of a larger effort aimed at imposing greater state control over the countryside.

A related component of this conservative retrenchment of central state authority over was the creation of the position of the rural land captain (*zemskii nachal'nik*) in 1889.[36] These officials, who were overwhelmingly members of the landed nobility, were nominated locally by assemblies of the nobility, approved by governors (or appointed by central authorities), and paid a salary by the Ministry of Internal Affairs. By the early twentieth century there were almost 2500 land captains in 43 provinces. Initially, these officials held almost complete authority over the townships governments and courts, including the possibility of imposing fines, imprisoning people, and even enacting corporal punishment. Over time, additional responsibilities were added, including administrative duties during the Stolypin land reforms of the early twentieth century. Answerable to governors and the Ministry of Internal Affairs, the land captains constituted an extension of central control into the countryside, but they held no fiscal authority of their own.

Critical to the work of the Offices of Peasant Affairs and the land captains were local police under the control of central authorities. Unlike deputies nominated and funded by the peasantry, the rural constables (*ispravniki*) and their staff were hired and paid by the Ministry. One of their main responsibilities was to help execute the allocation of external tax obligations among peasant communities, industrial establishments, and other local paying units. This occurred under the direction of district offices of the Ministry of Finance, the *zemstvo*, and municipal authorities. In this capacity, these police were also responsible for collecting tax arrears, which occasionally necessitated the forced sales of taxpayer assets and other punishments.

---

[35] See *PSZ*, Series II, vol. 39, no. 40457, clauses 90–91 and 94–98; and clause 87 of the 1890 reform law (Pearson 1989; Zakharova 1968).

[36] This section draws on Macey (1989) and Pearson (1989). Macey argues that the land captains reflected the growing "bureaucratization" of the countryside.

As hinted at above, reforms over the late Imperial period revised the degree to which outside authorities intervened in programs and budgets of the *zemstvo* and other institutions of local self-government. This included changes in the rules governing local levies and measures that shifted obligations for particular public services between different components and levels of the broader public sector. Overall, it is clear that the central government took an increasingly active interest in the provision of local public goods, especially after 1890, corresponding to the large increase in the size of the central government's budget in the last three decades of the Tsarist era (see Fig. 1).[37] Roadwork and other infrastructure improvements (including many private railroads) were eventually taken over by the state, as was troop quartering and the coordination of much local food relief efforts (particularly after the famine of 1891–1892). The Ministry of Education took an increasingly active role in supervising the system of primary education, culminating in a 1908 law that committed the state to the idea of universal primary education.[38] Similar central government efforts at local economic development emerged in the last three decades of the Imperial era in other areas, from agronomy to public health, but these were relatively limited and never completed abrogated the role of local actors capitalizing on local state capacity.[39]

Further Perspectives on Local State Capacity in Tsarist Russia

The previous sections have documented a complex structure of local governance in late-Tsarist Russia. Compared to the pre-1861 period, the formalization of the peasant institutions, coupled with the creation of the *zemstvo* and reforms of urban governance, led to a considerable increase in locally produced state capacity. The ability of different institutions to collect revenues increased substantially, which led to an increase in the provision of various public goods and services. Eventually, the central government acted to develop a greater presence at the local level, although this occurred relatively late and may have substituted, in part, for what was already occurring locally.

To give an idea of scale, total *zemstvo* (provincial + district) spending rose from approximately 4.4 % of central government spending in 1874 to just over 8 % in 1913 (Nafziger 2011). These expenditures were increasingly concentrated in health

[37] Amid perceptions of a growing rural economic crisis following the famine of 1891–1892, the central ministries viewed many *zemstva* as fiscally insolvent and began intervening more directly (Fallows 1982, pp. 216–217).

[38] See Eklof (1986). This measure required district *zemstva* to submit plans for achieving universal enrollment in their jurisdictions plans. In return, they received various subsidies and loans from the Ministry of Education. Growing state intervention in local educational matters also came in a succession of ministerial decrees and reforms from 1867, where the Ministry of Education took over supervision but not the funding of schools, to the 1908 law.

[39] On agronomy, see Nafziger (2013). On public health, see Frieden (1982).

care and education.[40] Add to that the rising capacities of the peasant institutions and urban governments, and it appears that much of the process of "state formation" was taking place outside of the orbit of the Petersburg Ministries and their local offices. Of course, this implies nothing about the efficiency by which revenue, legal, and infrastructural capacities were being built at any level. Indeed, the central government remained reliant upon less efficient indirect taxes that required less administrative capacity.

Regarding direct taxes, the state, peasant institutions, and the *zemstvo* may have competed for the loyalties and the tax dollars of the rural population as this process of local capacity building occurred.[41] The different levels government had a common primary tax base –assessments made on land and other forms of "immovable property." In terms of rural land, the state held about 110 million, peasant rural societies 120 million, and private owners roughly 100 million *desiatiny* (2.7 acres) subject to tax assessments in European Russia in 1905 (Loganov 1906). Urban properties and fixed capital constituted the other main direct tax bases. By 1905, the *zemstvo*, municipal governments, and local agents of state ministries had made considerable investments to determine the income-generating possibilities of different types of property held by different owners in different locations (Kotsonis 2014). At the bottom level, the multi-layered tax system relied on the principle of collective liability among peasant institutions and urban property owners. Although rates imposed by central and local authorities were set separately, the underlying base was valued in common, while collection was undertaken simultaneously as one assessment. As such, we can ask whether levies made by central authorities took priority in the building "capacity" in this context.

To examine this, we turn to data on the land-related obligations faced by the peasantry of European Russia in 1895.[42] On average, such peasant households held approximately 10 *desiatina* of land (c. 1905). With this in mind, Table 7 reports information on mean total land taxes per *desiatina*, as well as a breakdown by the type of assessment. Because we do not have exact information about the burdened

[40]Between 1885 and 1913, central government spending on education and health care rose from 23 million to 154 million rubles, or 2.7 to 4.6 % of total spending. Military spending stayed relatively constant at 27–29 % of overall expenditures throughout the period (Gregory 1982, p. 256). According to Eklof's tabulations (1986, p. 91), central government spending on primary education rose from only 0.3 % of the budget in 1862 to 2.225 % (or approximately 76 million rubles) in 1913. By 1913, *zemstva* spending on education – mostly primary – reached approximately 88 million rubles (Russia, Statisticheskoe 1913 vol.). Eklof (1986, p. 89) shows that central government contributions to rural primary schooling rose from 11.3 % of all funding in 1879 to 45 % in 1910, while *zemstvo* support fell from 43.4 % to 29.6 % over the same period. Some of these contributions took the form of subsidies and loans to *zemstva* to supplement existing or planned programs.

[41]Atkinson (1982), among others, notes the presence of fiscal and political conflict between the townships/communes and the *zemstva*.

[42]Other years and sources of similar data generate similar conclusions. Our focus in on peasant tax obligations related to land, as such detailed (district-level) data are unavailable for other classes and types of direct taxes.

Table 7 Peasant obligations and arrears for land-based direct taxes, 1895

	N	Mean	SD	Mean accumulated arrears/1895 obligations (×100)	SD
Total land-based taxes per *desiatina*	498	1.409	0.806	76.36	100
Shares of total direct taxes for:					
Central state obligations	499	57.89	10.92	101.53	146.2
Zemstvo	499	13.44	6.54	76.56	81
Township	499	11.57	8.82	21.31	22.22
Rural society	498	10.1	7.79	7.25	11.6
Obligatory fire insurance	499	7.08	4.22	n/a	n/a

Taxes per *desiatina* are in rubles. The *source* of the data is Russia, Department (1902)

population, we focus on the taxes per *desiatina*, although this obviously conflates land quality and tax burdens. The overall level of obligations was about 1.4 rubles per *desiatina*, with the majority owed to the central government, followed by the *zemstvo* and then the peasant institutions. The map depicted in the top of Fig. 5 shows two areas – right-bank Ukraine and the central agricultural provinces – that possessed considerably higher levels of land obligations.

Arrears on these land taxes accumulated to roughly 76 % of the 1895 assessment, which we view as quite low given the roughly 30 years over which this occurred. The bottom panel of Fig. 5 indicates that the districts along the northern Volga and to the northwest of Moscow had accumulated greater tax arrears by 1895. What is most telling is that the level of such non-payments appears to have been greater for the central government's demands and smallest for the rural societies. This suggests that local monitoring and the possibility or perception that closer levels of government provided more beneficial spending may have driven peasant payment strategies. While this evidence is certainly not definitive, it does illustrate the importance of considering the fiscal structures and capabilities of *both* central and local authorities in the Imperial Russian context.

Comparative and Conceptual Dimensions

Was the Russian case anomalous for this period? What was the nature of local state capacity building in other countries of Europe? The rich data that we have explored in this chapter are not easily accessible for other countries. Standard sources of fiscal information, such as Mitchell (1998), stick to central government revenues. Furthermore, important accounts of the emergence of the franchise and the modern social welfare state also emphasize the corresponding fiscal policies of *central* governments (e.g., Aidt et al. 2006; Dincecco 2011; Lindert 2004; Lizzeri and Persico 2004). A few recent works have delved into specific components of local government activity in other societies, but for the most part, these studies do not

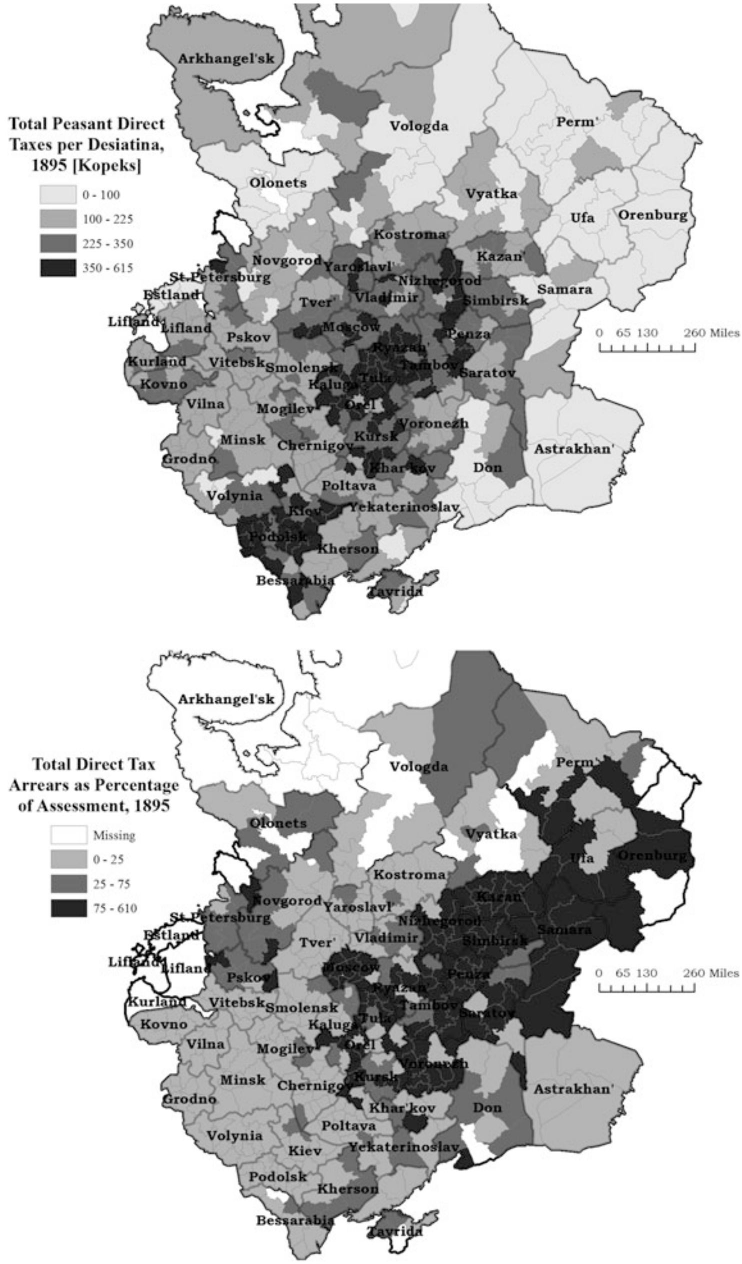


Fig. 5 Total direct tax obligations and arrears among the peasantry, 1895. Note: Figures drawn from data presented in Nafziger (2014). The amount of "obligated land" was defined in the original source. Payment arrears are defined as the accumulated amount relative to the year's assessment for 1895

emphasize the diversity of local entities, nor do they discuss the fiscal interactions between the center and locality.[43]

Moving slightly farther afield, recent empirical studies have investigated how the outcomes of decentralizing reforms in developing economies may be affected by the structure of local political institutions. Increasing the political voice of previously underrepresented groups (such as women, ethnic minorities, or lower social classes) can have significant effects on the amount and allocation of local public spending.[44] The impact of a decentralizing reform depends crucially on how an increase in the nominal political voice of a group is translated into real political influence through local institutions. In the Russian case, the newly empowered peasant institutions of self-government and municipal bodies were forms of decentralization, as was the creation of the *zemstvo*. Furthermore, achieving just minority positions by the peasantry in the *zemstvo* assemblies appears to have fostered opportunities to propose policies, obtain agenda-setting executive positions, and ally with elements of the other property owners to push through spending proposals (Nafziger 2011). All of these measures likely brought local spending and revenues more in line with the preferences of local residents, while possibly improving local accountability and leading to an increase in total public sector activity.[45]

Examining the Russian case suggests a modified approach towards the development of state capacity over the long nineteenth century. Rather than just focus on the political economy issues at the level of the central government, it is necessary to consider the entirety of the edifice that comprised the state from the top on down. Specifically, a fuller examination of the state's role in the transition to modern economic growth would consider spending policies of local governments. This naturally requires some consideration of the interrelationship between the center and the locality, whether in the form of fiscal federalism (Oates 1993), a decentralization or devolution of fiscal authority down the hierarchy (Bardhan and Mookherjee 2006), a structure of inter/intra-governmental grants, or in a strategic framework (e.g. Acemoğlu et al. 2014).[46]

[43]In Cardoso and Lains, eds. (2013), many of the chapters do acknowledge some complexity in the fiscal hierarchies of the nineteenth-century state, but they do not draw on the "capacity" framework. Economic history works that explicitly focus on one or two parts of local government include Chapman (2015), Legler et al. (1988), and Ziblatt (2008).

[44]Foundational contributions to what is a growing literature include Besley et al. (2004), Chattopadhyay and Duflo (2004), and Pande (2003), all on India.

[45]Although this presumes that local elites would not coopt these institutions more than national elite could capture more centralized revenue and spending policies, a point emphasized by Bardhan and Mookerjee (2006).

[46]Due to fears of coordinated opposition, the Tsarist state put explicit limitations on interactions between *zemstvo*; forbidding, for example, coordination in public health provision. Moreover, it was only with the onset of World War I that a serious discussion of an Empire-wide *zemstvo* system took place, with such a structure implemented in a limited way in 1917. Given this (and similar limits on other cross-border governance), the network model proposed by Acemoğlu et al. (2014), whereby local and central governments make their own capacity investments in a strategic way, is perhaps not entirely applicable in the Russian context.

Concluding Thoughts

Mark Harrison has constructed an influential narrative about how military strategy, armament production, and repression were critical drivers of the modern state's role in the economy, particularly when it came to developing the capacity to extract resources from the population. While Mark's work has focused on the Soviet period, military spending was an enormous part of the central government budget in Imperial Russia over the long nineteenth century, and external conflict and the threat of internal unrest surely did shape the development of the central authorities' fiscal and legal capacities (Pinter 1984). This paper argues that an exclusive focus on state formation as entirely a project of the central authorities misses an important local component, both in the Russian context and in European history, more generally.

In the case of Tsarist Russia, reforms in the wake of serf emancipation decentralized important fiscal and legal rights to a number of local institutions. This allowed for a significant amount of state capacity to be developed locally, rather than the central government simply imposing its policies and tax obligations. In providing some illustrative empirical evidence on just what these processes entailed, we do not want to oversell this point – as Tables 1 and 7 implied, central state direct and indirect obligations were considerably larger than those imposed by the local governments. However, much of this went to fund military and coercive aspects of the absolutist regime, rather than towards economic goals. It was left to the peasant institutions, municipal governments, corporate entities, and the new *zemstvo* to construct the capacity for most local public goods and services. While other scholars have begun to examine similar developments in Europe and the United States in the long nineteenth century, there is a large need to collect more and comparative information to better understand whether this locally produced capacity was a robust phenomenon. Nevertheless, bringing the *local* into the story of the rise of the modern nation state is an important task for economic historians, one that I am sure Mark Harrison would appreciate as well.

Archival and Published Primary Sources

Central Historical Archive of Moscow – TsIAM
Central State Historical Archive of St. Petersburg – TsGIA SPb
State Archive of Nizhegorod Oblast' – GANO
Russian State Historical Archive – RGIA

References

Acemoğlu, Daron, Camilo Garcia-Jimeno, and James A. Robinson. 2014. *State capacity and economic development: A network approach*. Working paper, MIT, Department of Economics.
Aidt, T.S., Jayasri Dutta, and Elena Loukoianova. 2006. Democracy comes to Europe: Franchise extension and fiscal outcomes 1830–1938. *European Economic Review* 50: 249–283.

Atkinson, Dorothy. 1982. The Zemstvo and the peasantry. In *The Zemstvo in Russia: An experiment in local self-government*, ed. Terence Emmons and Wayne S. Vucinich, 79–132. Cambridge: Cambridge University Press.

Bardhan, Pranab, and Dilip Mookherjee. 2006. Decentralisation and accountability in infrastructure delivery in developing Countries. *The Economic Journal* 116: 101–127.

Besley, Timothy, and Torsten Persson. 2010. State capacity, conflict, and development. *Econometrica* 78(1): 1–34.

Besley, Timothy, Rohini Pande, Lupin Rahman, and Vijayendra Rao. 2004. The politics of public good provision: Evidence from Indian local governments. *Journal of the European Economic Association* 2(2–3): 416–426.

Brewer, John. 1989. *The sinews of power: War, money, and the English state, 1688–1783*. New York: Knopf.

Brower, Daniel R. 1990. *The Russian city between tradition and modernity, 1850–1900*. Berkeley: University of California Press.

Cardoso, Jose Luis, and Pedro Lains, eds. 2013. *Paying for the liberal state: The rise of public finance in nineteenth-century Europe*. Cambridge: Cambridge University Press.

Chapman, Jonathan. 2015. The franchise, taxes, and public goods: The political economy of infrastructure investment in nineteenth century England. Working paper.

Chattopadhyay, Raghabendra, and Esther Duflo. 2004. Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica* 72(5): 1409–1443.

Dennison, Tracy K. 2011. *The institutional framework of Russian Serfdom*. Cambridge: Cambridge University Press.

Dincecco, Mark. 2011. *Political transformations and public finances: Europe, 1650–1913*. Cambridge: Cambridge University Press.

Dincecco, Mark, and Gabrial Katz. 2016. State capacity and long-run economic performance. *The Economic Journal* 126(590): 189–218.

Easley, Roxanne. 2008. *The emancipation of the serfs in Russia: Peace arbitrators and the development of civil society*. London: Routledge.

Eklof, Benjamin. 1986. *Russian peasant schools: Officialdom, village culture, and popular pedagogy, 1861–1914*. Berkeley: University of California Press.

Fallows, Thomas. 1982. The *Zemstvo* and the bureaucracy, 1890–1914. In *The Zemstvo in Russia: An experiment in local self-government*, ed. Terence Emmons and Wayne S. Vucinich, 177–241. Cambridge: Cambridge University Press.

Frieden, Nancy M. 1982. The politics of *Zemstvo* medicine. In *The Zemstvo in Russia: An experiment in self-government*, ed. Terence Emmons and Wayne S. Vucinich, 315–342. Cambridge: Cambridge University Press.

Gatrell, Peter, and Mark Harrison. 1993. The Russian and soviet economies in two world wars. *Economic History Review* 46(3): 425–452.

Gaudin, Corinne. 2007. *Ruling peasants: Village and state in late imperial Russia*. DeKalb: Northern Illinois University Press.

Gennaioli, Nicola, and Hans-Joachim Voth. 2015. State capacity and military conflict. *Review of Economic Studies* 82: 1409–1448.

Gregory, Paul R. 1982. *Russian national income, 1885–1913*. Cambridge: Cambridge University Press.

Gregory, Paul R., and Mark Harrison. 2005. Allocation under dictatorship: Research in Stalin's achieves. *Journal of Economic Literature* XLIII: 721–761.

Hamburg, Gary. 1984. *Politics of the Russian nobility, 1881–1905*. New Brunswick: Rutgers University Press.

Harrison, Mark (ed.). 1998. *The economics of world war II: Six great powers in international comparison*. Cambridge: Cambridge University Press.

Harrison, Mark, and Paul Gregory.

Hartley, Janet. 2006. Provincial and local government. In *The Cambridge history of Russia, volume 3: Imperial Russia, 1689–1917*, ed. Dominic Lieven, 449–467. Cambridge: Cambridge University Press.

Hoffman, Philip, and Jean-Laurent Rosenthal. 1997. The political economy of warfare and taxation in early modern Europe: Historical lessons for economic development. In *The frontiers of new institutional economics*, ed. John Drobak and John Nye, 31–55. San Diego: Harcourt Brace.

Ivanov, L.M. 1945. Gosudarstvennye krest'iane Moskovskoi gub. i reforma Kiseleva. *Istoricheskie Zapiski* 17: 76–129.

Koshman, L.V. 2008. *Gorod i gorodskaia zhizn' v Rossii XIX stoletiia*. Moscow: ROSSPEN.

Kotsonis, Yanni. 2014. *States of obligation: Taxes and citizenship in the Russian empire and early Soviet Republic*. Toronto: University of Toronto Press.

Lapteva, L. E. 1998. *Regional'noe i mestnoe upravlenie v Rossii (vtoraia polovina XIX veka).* Moscow: In-t gosudarstva i prava RAN.

Legler, John, Richard Sylla, and John Wallis. 1988. U.S. City finances and the growth of government, 1850–1902. *The Journal of Economic History* 48(2): 347–356.

Lindert, Peter H. 2004. *Growing public: Social spending and economic growth since the eighteenth century*, vol. 2. New York: Cambridge University Press.

Lindert, Peter H., and Steven Nafziger. 2014. Russian inequality on the eve of revolution. *The Journal of Economic History* 74(3): 767–798.

Lizzeri, Alessandro, and Nicola Persico. 2004. Why did the elites extend the suffrage? Democracy and the scope of government with an application to Britain's 'Age of Reform'. *The Quarterly Journal of Economics* 119: 707–765.

Loganov, G. 1906. *Statistika zemlevladeniia Evropeiskoi Rossii po uezdam*. St. Petersburg: Tip. Ministerstva vnutrennykh.

Macey, David A.J. 1989. The land captains: A note on their social composition, 1889–1913. *Russian History* 16(2–4): 327–351.

Mironov, Boris. 1985. The Russian peasant commune after the reforms of the 1860s. *Slavic Review* 44(3): 438–467.

Mitchell, B.R. 1998. *Europe, 1750–1993. Vol. 4 of international historical statistics*. London: Macmillan Reference Ltd.

Nafziger, Steven. 2011. Did Ivan's vote matter? The political economy of local democracy in Tsarist Russia. *European Review of Economic History* 15: 393–441.

Nafziger, Steven. 2013. Russian peasants and politicians: The political economy of local Agricultural Support in Nizhnii Novgorod Province, 1864–1914. In *Agricultural transformation in a global history perspective*, ed. Ellen Hillbom and Patrick Svensson, 108–135. New York: Routledge.

Nafziger, Steven. 2014. Serf emancipation, land reform, and factor endowments in Imperial Russia. Working paper.

Oates, Wallace. 1993. Fiscal decentralization and economic development. *National Tax Journal* 46(2): 237–243.

Pande, Rohini. 2003. Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. *The American Economic Review* 93(4): 1131–1152.

Pearson, Thomas S. 1989. *Russian officialdom in crisis: Autocracy and local self-government, 1861–1900*. Cambridge: Cambridge University Press.

Pinter, Walter. 1984. The burden of defense in imperial Russia, 1725–1914. *Russian Review* 43(3): 231–259.

Polnoe sobranie zakonov Rossiskoi Imperii [PSZ]. Multiple volumes. St. Petersburg, Russia.

Rieber, Alfred. 1991. *Merchants and entrepreneurs in imperial Russia*. Chapel Hill: The University of North Carolina Press.

Robbins, Richard G. 1987. *The Tsar's viceroys: Russian provincial governors in the last years of the empire*. Ithaca: Cornell University Press.

Rozhdestvenskii, S.V. ed. 1902. *Istoricheskii obzor deiatel'nosti Ministerstva narodago prosveshcheniia, 1802–1902*. St. Petersburg, Russia.

Russia. Departament Okladnykh Sborov, Ministerstvo finansov. 1917. *Dokhody, raskhody, spetsial'nye kapitaly i zadolzhennost' gorodskikh poselenii na 1912 god*. Petrograd, Russia.

Russia. Departament Okladnykh Sborov, Ministerstvo finansov. 1902. *Svod svedenii o postuplenii i vzimanii kazennykh, zemskikh i obshchestvennykh okladnykh sborov za 1895–99 g.g.* St. Petersburg, Russia

Russia. Ministerstvo finansov. 1900–1915. *Ezhegodnik Ministerstva Finansov.* Multiple volumes. St. Petersburg, Russia.

Russia. Ministerstvo vnutrennykh del. 1881. Svedeniia o sostav politseiskikh stanov i lichnom sostave uezdnoi politsii. St. Petersburg.

Russia. Statisticheskoe otdelenie, Departament okladnykh sborov, Ministerstvo finansov. 1908–1915. *Dokhody i raskhody zemstv 34-kh gubernii po smetam na 1903 [, 1906, and 1913].* Multiple vols. St. Petersburg, Russia.

Russia. Tsentral'nyi statisticheskii komitet, Ministerstvo vnutrennykh del. 1886. *Dokhody i raskhody gubernskikh i uezdnykh zemstv za 1883 God*. Series III, no. 16, *Statisticheskii vremennik Rossiiskoi imperii*. St. Petersburg, Russia.

Russia. Tsentral'nyi statisticheskii komitet, Ministerstvo vnutrennykh del. 1886. *Mirskie raskhody krest'ian za 1881 god*. Series III, no. 13. *Statisticheskii vremennik Rossiiskoi imperii*. St. Petersburg, Russia.

Russia. Tsentral'nyi statisticheskii komitet, Ministerstvo vnutrennykh del. 1896. *Zemskie raskhody, naznachennye po smetam i deistvitel'no proizvedennye v 1896 godu*. St. Petersburg, Russia.

Shakibi, Zhand P. 2006. Central government. In *The Cambridge history of Russia, volume 3: Imperial Russia, 1689–1917*, ed. Dominic Lieven, 429–448. Cambridge: Cambridge University Press.

Skalon, V. Iu (ed.). 1888. *Zemskiia dokhodnyia smety i raskladki na 1885 god*. St. Petersburg: Departament okladnykh sborov, Ministervo finansov.

Starr, S. Frederick. 1972. *Decentralization and self-government in Russia, 1830–1870*. Princeton: Princeton University Press.

Syrnev, A., ed. 1888. *Statistika vyborov v zemskiia uchrezhdeniia za 1883–1886 gg.* Vol. V, *Statistika Rossiiskoi Imperii*. St. Petersburg.

Veselovskii, Boris. 1909–1911. *Istoriia zemstva za sorok let*. 4 vols. St. Petersburg: O.H. Popova.

Waldron, Peter. 2007. *Governing Tsarist Russia*. London: Palgrave Macmillan.

Wirtschafter, Elise. 1997. *Social identity in imperial Russia*. DeKalb: Northern Illinois University Press.

Yaney, George. 1973. *The systemization of Russian government*. Urbana: The University of Illinois Press.

Yun-Casalilla, Bartolome, and Patrick O'Brien (with Francisco Comin Comin), eds. 2012. *The rise of fiscal states: A global history, 1500–1914*. Cambridge, UK.

Zakharov, V.N., Iu A. Petrov, and M.K. Shatsillo. 2006. *Istoriia nalogov v Rossii, IX – nachalo XX veka*. Moscow: Rosspen.

Zakharova, Larissa. 1968. *Zemskaia kontrreforma 1890 goda*. Moscow

Zhurnaly Ardatovskago uezdnago zemskago sobraniia Simbirskoi gubernii. Ocherednoi sessii 1883 g. i ekstrennykh sozyvov na 11 Ianvaria i 12 Iunia 1883 goda. Nizhii Novgorod: Roiskii and Dushin. 1884

Ziblatt, Daniel. 2008. Why some cities provide more public goods than others: A subnational comparison of the provision of public goods in German cities in 1912. *Studies in Comparative International Development* 43: 273–289.

Historicizing Divergence: A Comparative Analysis of the Revolutionary Crises in Russia and Finland

Pavel Osinsky and Jari Eloranta

Abstract Beginning 1917, Russia and Finland both experienced revolutionary situations, the seizure of power by radical political groups, and civil wars. However, the ultimate outcomes of the revolutionary crises in the two countries turned out to be different: the Russian Bolsheviks won the struggle for power whereas the Finnish Red Guard suffered a defeat. Why did the radical socialists win in Russia but lose in Finland? This chapter argues that the Russian revolutionaries benefited from the existence of two coalition alliances that had not fully materialized in Finland: the workers–soldiers' alliance, which was critical for the radicals' seizure of power, and the workers–peasants' alliance, which became pivotal during the years of the civil war. Thus, our comparative historical analysis lends support to the "social history" of the revolutions but—in contrast to other writings—draws attention to the centrality of structural conditions created by a mass mobilization war and the contingent nature of the extant revolutionary alliances.

Keywords First World War • Russia • Finland • Civil war • Revolution • Alliance

Introduction

Few events in twentieth-century history attracted as much scholarly attention as the Russian Revolution. Monographs, articles, and memoirs examining this momentous event are countless. Most of these narratives are based on the assumption that this transformation was prepared by the whole course of Russia's development. As it

P. Osinsky
Department of Sociology, Appalachian State University, 209 Chapell Wilson Hall, Boone, NC 28608, USA
e-mail: osinskyp@appstate.edu

J. Eloranta (✉)
Department of History, Appalachian State University, 224 Joyce Lawrence Lane, Boone, NC 28608, USA

Department of History, University of Jyvaskyla, Finland
e-mail: elorantaj@appstate.edu

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_4

follows from such accounts, the 1917 revolution was virtually inevitable. However, such reasoning neglects the fact that a number of countries at a comparable level of socioeconomic development and placed in similar political conditions, including the former provinces of the Russian empire, did not experience such transformation or experienced it in a drastically different way.

The political turmoil in Finland in 1917–1918 represents a particularly interesting case from that standpoint. Being a part of the Russian empire until the declaration of independence in December of 1917, Finland was involved in the whirlpool of the events initiated by the Russian Revolution. The Finnish radicals launched a frontal attack on the institutions of the old order. The Finnish Civil War, which broke out in January of 1918, represented the first of Europe's revolutionary civil wars of the twentieth century (Payne 2011).[1] Nonetheless, the Finnish radicals had failed. In the civil war, which lasted from late January until early May of 1918, the Finnish White Army organized by General Mannerheim defeated the Red Finns. The revolutionary movement was suppressed amidst violence and terror. How can one explain such divergence of the revolutionary outcomes in Russia and Finland given that many of the initial conditions were similar?

Like many comparative analyses, this research does not examine primary sources but rather summarizes the existing findings and highlights the patterns of similarity and dissimilarity in a paired comparison (for paired comparison methodology see Lange 2013; Slater and Ziblatt 2013; Tarrow 2010). In the next section we will discuss the debates about these revolutions, to be followed by deeper analyses of the Russian and Finnish cases.

Working Class Revolutions?

To understand why a Russian-type revolution failed in Finland, one needs to first identify the nature of the Russian revolution. Historians and historical sociologists have formulated several interpretations of it (see a review in Kolonitskii 2009). One school of thought insisted on a political interpretation of the Russian revolution, claiming that it was primarily a product of a purposeful action of the radical socialist party, the Bolsheviks. This tradition emphasized the role of the leadership, organization, and ideology as the key factors of the Bolshevik victory (e.g., Jowitt 1992; Malia 1994; Pipes 1990). Another, a revisionist school of historians, argued that this transformation was prepared by the whole of Russia's socioeconomic development and driven by large segments of the population, including the working class, the poor and middle peasants, a part of the educated class, the ethnic minorities, etc. (e.g., Figes 1996; Riga 2012).

The latter, structural interpretations of the revolution can be divided into two subcategories. Some early studies emphasized the peculiarity of Russia's agricultural class structure and viewed the cataclysm of 1917 as a peasant-based revolution

[1] The full-scale military action of the Russian civil war began in May of 1918.

(e.g., Moore 1966; Skocpol 1979). Other contributions have interpreted the Russian revolution as a predominantly working class movement (e.g., Koenker and Rosenberg 1989; McDaniel 1988). The boundaries between various interpretations remained blurred and many accounts of the revolution described the course of events without giving preference to either version (e.g., Badcock 2007; Wade 2005). It was accepted, however, that a revisionist interpretation became a dominant perspective and that the conception of the working-class revolution grasped the crux of the event better than any other narrative (Suny 1994).

The intriguing part of the story is that the political process in Finland closely followed the Russian events and represented, to a great extent, part and parcel of the Russian revolution.[2] The insurrection in January 1918 and the subsequent civil war may be seen as a genuine working class revolution. The Red Finns, like their Russian counterparts, enjoyed important strategic advantages in the beginning of the civil war. In February and March of 1918 they controlled the economically advanced and densely populated south of Finland where Helsinki and other major cities were located. They commanded railroads and lines of communication. They obtained a large number of weapons and military supplies from the Bolshevik government in Petrograd as well as the Russian troops quartered in Finland. Nonetheless, it was the White army that eventually won the civil war. So, once again, why Russia and not Finland?

What Was Different?

This study is premised on a war-centered approach in contemporary social science which views warfare as the primary driving force of state formation (e.g., Downing 1992; Gat 2006; Glete 2001; Porter 1994; Tilly 1990). As Charles Tilly (1975: 42) had famously stated, "war made the state, and the state made war." From such a point of view, existing interpretations of communist revolutions share one common limitation: they fail to place these events in the context of the mass mobilization wars that erupted in the first half of the twentieth century. Surely, most accounts of the communist revolution in Russia duly stress the role of war in undermining the old regime, along, however, with numerous other factors. Our argument goes beyond just highlighting the impact of war or the military power. It suggests that war was a primary cause, condition, and context of the paradigmatic communist revolutions. To put it simply: no mass mobilization war, no communist revolution, especially in a large country (Osinsky and Eloranta 2014).

Proceeding from this thesis we argue that divergence in experiences of the two countries in the First World War was the most important factor behind the bifurcated

[2]For historiography of the Finnish civil war, see Tepora and Roselius (2014).

dynamics of political conflicts in Russia and Finland.[3] Although many European nations confronted social unrest by the end of the war, in 1917 Russia's economic collapse and political crisis resulted in the forming of the revolutionary alliances that did not emerge elsewhere. The first such alliance was a coalition of urban workers, who started anti-government protests, and the rank-and-file soldiers, garrisoned in the cities, who took the side of the workers. Petrograd workers and soldiers overthrew the autocracy in February of 1917 and sustained the revolutionary process in the following months. Their alliance was also pivotal during the Bolsheviks' takeover in October and the dissolution of the Constituent Assembly in January 1918. The second alliance, which became critically important in the next phase, the civil war, was a coalition of the urban-based radicals and the peasantry. In October of 1917, Lenin legitimized the peasants' claims on the land and thus laid a foundation of the Bolsheviks' alliance with the peasants. Although subsequent relations between the new regime and the rural population were uneasy, in the critical period of the civil war most peasants chose to align with the Reds rather than the Whites.

In Finland, where the population did not directly participate in the war, a military–civilian alliance against the authorities failed to materialize. Finns did not have their own army and the navy that would provide fighting crusaders for the revolutions that happened in Russia. The revolutionary action was carried out almost exclusively by the radical segments of the working class organized in the Red Guard, which were formed rather spontaneously and lacked a centralized structure. The urban–rural alliance also turned out to be quite tenuous. A land problem was not as acute in Finland as in Russia. Such a powerful source of rural radicalism as the gentry's landownership was not much of an issue in Finland, where the independent landed nobility had almost disappeared. In addition, the farming population was sharply divided along class lines: the landless laborers and tenants were inclined to support the urban revolution, yet the independent smallholders opposed it. In the rest of the chapter we elaborate on these propositions.

The Military–Civilian Alliance in Russia

The Russian revolution of 1917 cannot be understood outside the context of Russia's experience in the First World War. Not only did the First World War cause major economic and social dislocations, it also shook the core institutions of the Russian state, including the army. To fight the enemy troops, the government drafted about

[3]The First World War resulted in unprecedented mobilization of human and material resources by the belligerent nations (Broadberry and Harrison 2005). The revolutions and civil wars, which broke out in the wake of the war, represented, as we have shown elsewhere, a product of disintegration and political realignment of multi-million armies and societies drafted to fight enemy states (see Osinsky 2008; Osinsky and Eloranta 2014).

15 million men, more than any other nation in the war. However, due to poor organization, inadequate equipment, and incompetent commanders, the Russian armies suffered defeats. To replace the battlefield losses (about two million dead and five million prisoners of war), the government had to mobilize new cohorts of recruits and reservists. Petrograd, Moscow, and other large cities had been turned into large military training facilities. In the meantime, due to disorganization of transportation and ineffective food supply policies, the urban population began to experience shortages of provisions and other living necessities. In the third winter of war (1916–1917), hundreds of thousands of people came out with protests against the war and the monarchy (see Gatrell 2005).

It is true that the revolutionary events in Petrograd in February of 1917 began with the strikes and rallies of the working class. These protests, however, did not bring about the fall of the government until the mutiny of soldiers of the Petrograd garrison on February 27. When the guard regiments, one after the other, turned to the side of the workers, the government resigned, and the power was transferred to the leaders of the State Duma and the Petrograd Soviet. The Romanov's monarchy, which had lost the support of the army, expired a few days later. An impromptu alliance of the labor movement and the insurgent soldiers laid the foundation of the plebeian–praetorian coalition (the "revolutionary democracy," in the language of the streets) that was to play a key role in subsequent events (Kolonitskii 2004).

In the following months, the main forces of the "revolutionary democracy"—the Kronstadt sailors, the Red Guards, and the Petrograd garrison—remained in the background of events, but every time the political situation turned serious, they appeared in the streets of Petrograd. The Provisional Government became a de facto hostage of the Petrograd garrison and the Red Guards. When the government decided to send a large part of the unruly garrison to the front in October, the soldiers decided to overthrow the Provisional Government. On October 24–25, the military units loyal to the Bolsheviks took control of Petrograd. In the next few days, the Soviet authority spread to other cities of Russia, where the outcome of the conflict was determined, in most cases, by the action of the soldiers garrisoned in these cities (Figes 1996; Frenkin 1982).

Thus, in a broad historical context of war and revolution, the statement that the Bolsheviks seized power in October is only partially accurate. A larger political force behind the Bolsheviks was the urban-based coalition of workers and soldiers, which would not emerge under conditions other than an unsuccessful experience in a mass-mobilization war and would not take over the state authority unless faced with the demoralization and defeat of the army.

The Military–Civilian Alliance in Finland

Being a semiautonomous province of the Russian Empire, Finland was not able to escape the experience of war completely. The Czarist government increased the number of Russian troops quartered in Finland to 50,000 and Helsinki became the main operational base of the Baltic fleet. As the war dragged on, the economic conditions and the living standards of most of the urban population, particularly industrial workers, worsened, leading to riots and strikes. Nonetheless, if one would inquire whether or not a broad-based military–civilian opposition to the existing authorities emerged in Finland as it did in Russia, the answer would be negative. The Duchy of Finland did not have its own army. A draft was not conducted in Finland and the population did not suffer losses of human life. The country was not a war zone; it was neither invaded nor occupied by enemy troops. The effect of war on the economy, infrastructure, and urban society in Finland was not as disastrous as in Russia. Apart from the Russian military presence, Finland had many characteristics of a neutral country (Haapala 2014; Upton 1980).

The February revolution of 1917 in Russia invigorated the national independence movement and emboldened labor activists but did not disrupt the economic and political life of Finland. The main political institutions (the parliament, the government, the political parties, the local administration) and civil society remained largely intact. In contrast to post-February Russia, where war politics moved to the center stage and where soldiers and civilians participated in large antiwar rallies, most conflicts in Finland at the time revolved around local economic and social issues. The most visible change was the disappearance of the old police and emergence of the numerous local militias, such as the Red Guard, formed by the industrial workers, and the Civil Guard, organized by middle-class individuals. As a result of military buildup, Finnish society became increasingly polarized into two power blocs: the Social Democrats with the Red Guards on the one side, and the conservative government, supported by the Civil Guards on the other side (see Siltala 2014).

The uprising of the Red Guards in January of 1918, the first act of the Finnish Civil War, began when the momentum created by the Bolshevik uprising in Russia had been lost. By that time the bourgeois government of P.E. Svinhufvud had declared Finnish independence and consolidated its power. Until the end of January, the Civil Guards, formed for defending the middle class and their property, functioned as a self-governed citizens' militia. On January 25, however, they were officially declared the troops of the government. In response, on the night of January 28, the Red Guards took control of Helsinki. Government offices, telegraph, telephone, railway stations, banks, and other strategic points were occupied. The commander of the Red Guards, Eero Haapalainen, issued a manifesto proclaiming the Socialist Workers' Republic of Finland. The new revolutionary government, the People's Delegation, was formed, whereas the parliament and the old cabinet were dissolved (Alapuro 1988; Haapala 1993; Siltala 2014).

In the initial phase of the civil war, the Red Guards were able to keep the initiative. The Nationalist army, organized by General Mannerheim to fight the insurgents, was small and short on munitions. The presence of well-trained Russian troops, which helped the Red Finns with weapons and (occasionally) with direct military assistance, was still a major factor. In mid-March, however, the White troops were able to take over the initiative. Since the Great War was over, Russian troops began to withdraw from Finland. Mannerheim, on the other side, initiated conscription and ordered to "shoot on sight" people who resisted the White forces or engaged in sabotage. The battle for Tampere, the major industrial center of Finland, was the turning point of the war. In this battle the Reds suffered a defeat from which they would never be able to recover (Ylikangas 1993). The White army, supported by the German expeditionary troops that landed in early April, methodically expanded the territory under its control. In mid-April, Helsinki fell to the counterrevolutionary forces without a fight. By early May, all of Finland was cleared of the Red Guards (Kronlund 1989).

Why did the revolutionary government in Finland suffer defeat? Finland's participation in the war and the scope of the political disorganization it experienced were rather limited. In contrast to the Russian revolution, which erupted in a country devastated by war and was driven by large units of war-weary soldiers and sailors, the revolutionary process in Finland began when the national conscripted army did not exist whereas most political institutions in Finland remained intact, albeit divided by the competing factions. The parliament, the central government, the local administration, the public organizations, and the courts continued to operate. Many Finns abstained from participating in the conflict altogether. The insurgency was sustained almost exclusively by the working class militias, which lacked a centralized structure, possessed few military skills, and displayed low discipline (Manninen 1992).

The Urban–Rural Alliance in Russia

A similar divergence can be observed with respect to the urban–rural alliances. Before the revolution, Russia was an overwhelmingly rural country in which peasants made up about 80 % of the population. Although it had made some advances toward industrialization, modern institutions had little impact on the realities of a rural world culturally separated from urban society and governed by its own customs and traditions. The islands of estates of the nobility coexisted with a sea of peasant smallholdings managed through traditional communal arrangements that allowed adaptation to the changing natural and social conditions (Nafziger 2010). Most squires were absentee landlords who did not reside in their own estates. Because land was the primary factor of agricultural production, the property of gentry and well-to-do peasants (*kulaks*) represented the most desired possessions for less prosperous cultivators. Russia's countryside was in a state of recurrent rural unrest (e.g., Smith 2011).

For the Bolsheviks, who came to power in October, the peasants and their problems were of secondary importance. However, once political conflict turned to a critical phase, ignoring the concerns of the majority of the Russian people would have been a political mistake. One of the first documents adopted by the Lenin's government was the Decree on Land, which announced the abolition of private property on land (meaning squires' estates). It stipulated that all land was to be placed at the disposition of the district land committees and would be distributed equally among the peasants. This move made by the Bolsheviks provided a basis for an alliance between the new government and the peasants that played a critical role in the course of the civil war.

The first act of the Russian civil war opened with the mutiny of the Czechoslovak corps on the Trans-Siberian railroad in May of 1918. By banishing the Bolshevik authorities from the major cities on the Volga and along the Trans-Siberian railroad, the Czechoslovaks created a political vacuum soon to be filled by the anti-Bolshevik socialist opposition, which formed a new authority, the Committee of the Members of the Constituent Assembly, or the Komuch. This government presented itself as the legitimate successor of the Russian parliament. However, outside the circles of provincial intelligentsia, students, and cadets, the social constituency of this government remained small. Neither workers nor bourgeoisie trusted this middle-of-the-road authority. Most importantly, the Komuch failed to appeal to the rural population. The Volga peasants saw no reason to fight the Bolsheviks. The revolution gave them land and freedom and that was all that peasants really wanted. As a result, a draft campaign for a "People's Army" failed to produce enough soldiers. In the autumn of 1918, the Komuch troops were defeated by the Bolshevik forces whereas the Komuch politicians were driven to the Urals and Siberia (Kondrashin 2009; Mawdsley 2005).

The next year became the critical period of the civil war. The Bolsheviks had to fight three White armies in succession: the forces of Admiral Kolchak that advanced from Siberia, the army of General Denikin moving towards Moscow from Southern Russia, and General Yudenich's troops that were threatening Petrograd from the west. It would be misleading to characterize the relations between the peasants and the authorities as idyllic. The Bolshevik government used massive coercion to mobilize rural population and necessary resources. In many provinces, peasants rebelled against the draft and requisitions (Brovkin 1994; Osipova 2001; Raleigh 2002; Retish 2008).

In the context of these circumstances, the observation that most peasants disliked both protagonists ("a plague on both your houses") and preferred to keep out of the "war between brothers" (e.g., Pipes 1995: 272) is generally correct. Yet, it glosses over an important difference. However much the peasants disliked the Reds, their hatred of the Whites, who promised to restore the land to the landlords, ran far deeper (Kara-Murza 2003; Suny 2010). Even in areas where land was not the issue, the Whites emblematized the old regime, oppressive and odious (Lonergan 2008). True to their old habits, most White Guards treated persons of the "lower estates" with suspicion and contempt. The White troops assisted the former landlords in reclaiming their estates and punishing the peasants who seized their property (Figes 1996).

That explains a peculiar pattern in the Red Army recruitment. Through most of the civil war, the local recruitment committees enlisted peasants in large numbers, but the majority of the recruits deserted. When, however, Kolchak's troops launched an offensive against the Bolsheviks in the East, many deserters began returning to the Red Army. As a result, starting with 800,000 soldiers in January, the Red Army doubled in size by the end of April, the height of Kolchak's offensive in the east.

Most recruits came from the Volga region, where the peasants had made substantial land gains and had the most to fear from a White victory (Figes 1996).

The same change occurred a few months later, during Denikin's advance to Moscow from the south in the autumn of 1919 (Mawdsley 2005). The threat of a White victory made the peasants fear for the loss of their newly acquired land, and therefore they chose to return to the army (Figes 1990, 1996). By January 1920, the Red Army had three million men under arms and by the end of the year five million, whereas the combat effectiveness of the White armies never exceeded 250,000–300,000 and the number of all enlisted recruits one million (Galin 2006; Pipes 1993).

The Urban–Rural Alliance in Finland

Weak support of the Finnish urban-based revolution in the countryside was a major cause of the radicals' defeat. As stated above, the key factor of the Red Army's success in Russia was effective Bolshevik recruitment of millions of peasants. In the critical months of the civil war, Russia's farmers rose to defend their land and freedom, shifting the balance of military power to the side of the revolution. In Finland, however, the opposing armies turned out to be roughly comparable in size; at the peak of the conflict each side deployed about 70,000–80,000 troops. For the Reds, who advocated the people's interests, it was a poor performance. Why had the Red Finns failed to mobilize the rural masses in a way the Bolsheviks had done it in Russia?

To explain why most peasants remained indifferent or hostile to the urban revolution, we need to look at the class structure of Finland's rural population. Risto Alapuro (1988) identified three major groups of the Finnish agrarian population: the agricultural laborers, the crofters, and the freeholders. The first group was composed of the landless agricultural laborers who made up as much as 48 % of all rural households in the early twentieth century. These laborers were poor but did not experience truly acute economic distress in the early twentieth century. The rural commercial economy expanded, and wages increased (Ojala and Nummela 2006). Before the war, the living standards of the rural laborers improved substantially. The crofters, who made up about 17 % of the rural households, were the small leaseholders who paid rent to the landowners by working a certain number of days for the owner or making payments in cash or in kind. Some of the crofters joined the revolution, whereas others did not. Why? In Southern Finland, many tenants of large manors supported the Reds because they promised recognition of their property rights to land. In the North, many crofters worked on their relatives' land and defended their property rights against the socialists (Haapala 2014: 32). Finally, the independent peasant landowners, who made up about 35 % of all households, represented the most conservative part of the rural population. In the prewar period, landowners benefited from rapid commercialization of agriculture, particularly in such market-oriented industries as forestry and dairy husbandry. The peasant land possessions expanded. Many freeholders hired laborers for the harvest seasons or leased their land to tenants. Between 1870 and 1910, the net property of the peasants increased seven times (Arosalo 1998: 151).

How did these processes affect the military–political mobilization in the countryside in early 1918? Although both sides made efforts to mobilize supporters in the rural areas, the composition of their troops turned out to be quite different. The Red Guards consisted mostly of the urban and rural workers (Rasila 1969). Of the more than 3500 Reds killed at the front, 63 % were industrial workers, 16 % farm workers, 13 % tenants, 5 % farmers, 1 % civil servants, and 2 % others. The White Army, on the other side, consisted mainly of the recruits from the independent peasantry as well as the upper and the middle classes. Among the victims of the Reds, the farmers made up 45 %, civil servants 17 %, industrial workers 14 %, tenants 11 %, farm workers 9 %, and others 4 % (Manninen 1978: 233–239). Thus, the more affluent peasants and the middle classes formed the backbone of Mannerheim's troops.

Discussion and Conclusions

The end of the Cold War and the demise of the Communist regime in the Soviet Union renewed the discussions about the nature of the Russian and Finnish revolutions. The opening of previously closed archives allowed filling many gaps in the history of these events. The newly available data indicated, for instance, that support for the Bolshevik regime among Russia's lower classes was far from being universal. Historians discovered a significant amount of evidence documenting workers' revolts against the Bolsheviks' government (e.g., Brovkin 1990; Churakov 2007; Dmitriev and Kulikov 1992) and the veritable peasant war waged against the policies of the new regime in the countryside (e.g., Graziosi 1997; Kondrashin 2009; Osipova 2001; Pavlyuchenkov 1996). This new information challenged the "social history" interpretations of the Russian revolution, which subsequently lost much of their former academic impregnability. The politics-oriented "history from above" returned to the academic discourse with a vengeance (see Malia 1994).

This chapter provides new arguments in support of viewing the revolutions in Russia and Finland as structurally determined political conflicts involving large groups of the population. We need not a retreat from the social history of the revolutions but rather a change in the way we frame them. This change, we believe, requires reorienting our analyses in two ways. First, we need to place the revolutionary events in Russia, Finland, and other European countries in the context of the First World War (see Holquist 2002; Lieven 2015). Although scholars acknowledged a connection between the First World War and the rise of collectivist regimes long ago, they have not shown well enough how the wartime mass mobilization determined the political transformations of the European nations in the interwar period (see Porter 1994 for an exception). Second, this proposed reorientation implies focusing analyses of the revolutionary events around the notion of a coalition rather than simply a class. The revolution of 1917 in Russia was not just a proletarian revolution but a concatenation of several revolutions: workers vs. employers, soldiers vs. commanders, peasants vs. landlords, regionalists vs. centralizers, ethnic minorities vs. the ethnic majority, and so on. It was the coalitional nature of the revolutionary movement that ultimately made it triumphant.

The Finnish revolution, contrariwise, in many ways exemplified a class conflict (Kissane and Sitter 2005; Tepora and Roselius 2014). Because the country was not directly involved in a total war, the military component of the revolution was not as accentuated as in Russia. The revolution was driven by the radical workers organized into the Red Guard, a decentralized proletarian militia. Despite the fact that a large contingent of Russian troops was quartered in Finland and many Russians were sympathetic to the Red Finns' cause, their participation in the Finnish Civil War was limited. After the armistice with the Germans was signed, the Russian soldiers' radicalism weakened; the thing they wanted most was to go home. The mostly urban Red Finns were left to fight for their cause alone. Lacking centralized organization and military training, the workers' militias were not able to resist the better organized and trained army of General Mannerheim, which also secured military and logistical support from Germany.

The Red Army in Russia was initially destined to fare as poorly as the Red Finns until it acquired a centralized hierarchical structure and expanded its ranks. The Bolsheviks exerted great effort to create a functional command structure and bring in the "military specialists" (i.e., the former Czarist officers) to the Red army. By the time of the outbreak of the civil war, the new regime possessed institutions capable of conducting the conscription and training the troops (see Sanborn 2002). The transfer of the gentry's land to the peasants helped Lenin to ensure the peasants' neutrality in the conflict against the moderate socialists traditionally supported by the rural population. Although the Bolsheviks' requisitions and confiscations often provoked the hostility of peasants, they changed their attitude when the White armies launched the decisive campaigns in 1919. Confronted with the White Guard officers and the Cossacks who attempted to undo the peasants' revolution on land, many rural dwellers chose to go along with the Bolsheviks (also see Kara-Murza 2003).

Finnish urban-based revolutionaries, on the other hand, obtained only limited support in the countryside. Compared with Russia, where the peasantry was united in its opposition to the nobility and acted as a single group, such a source of rural radicalism as the gentry's land ownership was insignificant in Finland. Furthermore, the farmers, in contrast to Russian peasantry with its centuries-old egalitarian tradition, were sharply divided along class and property lines. Whereas many crofters and laborers welcomed the revolution and joined the ranks of the Red Guard, independent smallholders, who represented the wealthiest and the most respected part of rural communities, ultimately turned against the urban revolution. Many rural dwellers preferred social stability and opposed a violent armed conflict. As a result, it was Mannerheim's army rather than Red Finns who benefited more from the rural political support. As it follows from our analysis, both radical coalitions, the military–civilian and the urban–rural, were contingent and situational in their nature, and were crucial in determining the success of the Bolsheviks in Russia. These factors were not present in such a fashion in the Finnish case.

References

Alapuro, Risto. 1988. *State and revolution in Finland*. Berkeley: University of California Press.

Arosalo, Sirkka. 1998. Social conditions for political violence: Red and White Terror in the Finnish Civil War of 1918. *Journal of Peace Research* 35: 147–166.

Badcock, Sarah. 2007. *Politics and the people in revolutionary Russia: A provincial history*. Cambridge: Cambridge University Press.

Broadberry, Stephen, and Mark Harrison. 2005. The economics of World War I: An overview. In *The economics of World War I*, ed. S. Broadberry and M. Harrison, 1–40. Cambridge: Cambridge University Press.

Brovkin, Vladimir. 1990. Workers' unrest and the Bolshevik response in 1919. *Slavic Review* 43: 350–373.

Brovkin, Vladimir. 1994. *Behind the front lines of the civil war*. Princeton: Princeton University Press.

Churakov, Dimitry. 2007. *Buntuyushie proletarii: Rabochii protest v Sovetskoi Rossii* [The rebellious proletarians: workers' protest in the Soviet Russia]. Moskva: Veche.

Dmitriev, Pavel, and Kuz'ma Kulikov. 1992. *Myatezh v Izhevsko-Votkinskom raione* [The uprising in the Izhevsk-Votkinsk district]. Izhevsk: Udmurtiya.

Downing, Brian M. 1992. *The military revolution and political change: Origins of democracy and autocracy in early modern Europe*. Princeton: Princeton University Press.

Figes, Orlando. 1990. The Red Army and mass mobilization during the Russian Civil War. *Past & Present* 129: 168–211.

Figes, Orlando. 1996. *A people's tragedy: The Russian revolution, 1891–1924*. London: Penguin.

Frenkin, Mikhail. 1982. *Zahvat vlasti bolshevikami v Rossii i rol' tylovyh garnizonov armii* [The Bolsheviks' seizure of power in Russia and the role of the rear garrisons of the army]. Jerusalem: Stav.

Galin, Vasily. 2006. *Zapretnaya Politeconomiya: Krasnoe i Beloe* [The forbidden political economy: Red and white], vol. 2. Moskva: Algoritm.

Gat, Azar. 2006. *War in human civilization*. Oxford: Oxford University Press.

Gatrell, Peter. 2005. *Russia's First World War: A social and economic history*. Harlow: Pearson.

Glete, Jan. 2001. *War and the state in early modern Europe: Spain, the Dutch Republic, and Sweden as fiscal-military states*. New York: Routledge.

Graziosi, Andrea. 1997. *A great Soviet peasant war: Bolsheviks and peasants, 1917–1933*. Cambridge, MA: Harvard University Press.

Haapala, Pertti. 1993. Luokkasota. *Historiallinen Aikakauskirja* 2.

Haapala, Pertti. 2014. The expected and non-expected roots of chaos: Precondition of the Finnish Civil War. In *The Finnish Civil War: History, memory, legacy*, ed. T. Tepora and A. Roselius, 21–50. Leiden: Brill.

Holquist, Peter. 2002. *Making war, forging revolution: Russia's continuum of crisis, 1914–1921*. Cambridge, MA: Harvard University Press.

Jowitt, Kenneth. 1992. *The new world disorder: The Leninist extinction*. Berkeley: University of California Press.

Kara-Murza, Sergei. 2003. *Grazhdanskaia voina, 1918–1921: Uroki dlya XXI veka.* [The Civil War, 1918–1921: The lessons for the XXI century]. Moskva: Algoritm.

Kissane, Bill, and Nick Sitter. 2005. Civil wars, party politics and the consolidation of regimes in twentieth century Europe. *Democratization* 12: 183–201.

Koenker, Diane, and William Rosenberg. 1989. *Strikers and revolution in Russia, 1917*. Princeton: Princeton University Press.

Kolonitskii, Boris. 2004. "Democracy" in the political consciousness of the February revolution. In *Revolutionary Russia: New approaches*, ed. Rex A. Wade, 75–90. New York: Routledge.

Kolonitskii, Boris. 2009. Russian historiography of the 1917 revolution. *History & Memory* 21: 34–59.

Kondrashin, Viktor. 2009. *Krest'ianstvo Rossii v Grazhdanskoi voine.* [The peasantry of Russia in the Civil War]. Moskva: ROSSPEN.

Kronlund, Jan. 1989. *Suomen puolustuslaitos 1918–1939. Puolustusvoimien rauhan ajan historia, Sotatieteen laitoksen julkaisuja XXIV, Sotahistorian toimisto*. Porvoo: WSOY.

Lange, Matthew. 2013. *Comparative-historical methods*. Thousand Oaks: Sage.

Lieven, Dominic. 2015. *The end of Tsarist Russia: The march to World War I & Revolution*. New York: Viking.

Lonergan, Gayle. 2008. Resistance, support and the changing dynamics of the village in Kolchakia during the Russian Civil War. *Revolutionary Russia* 21: 57–72.

Malia, Martin. 1994. *The soviet tragedy: A history of socialism in Russia, 1917–1991*. New York: Free Press.

Manninen, Ohto. 1978. Red, white and blue in Finland, 1918. *Scandinavian Journal of History* 3: 229–249.

Manninen, Turo. 1992. *Järjestysvalta järkkyy: Kaartit vastakkain. In Itsenäistymisen vuodet* I: Irti Venäjästä. VAPK Kustannus.

Mawdsley, Evan. 2005. *The Russian Civil War*. New York: Pegasus Books.

McDaniel, Timothy. 1988. *Autocracy, capitalism, and revolution in Russia*. Berkeley: University of California Press.

Moore Jr., Barrington. 1966. *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.

Nafziger, Steven. 2010. Peasant communes and factor markets in late nineteenth-century Russia. *Explorations in Economic History* 47(4): 381–402.

Ojala, Jari, and Ilkka Nummela. 2006. Feeding economic growth: Agriculture. In *The road to prosperity: An economic history of Finland*, ed. J. Ojala, J. Eloranta, and J. Jalava, 65–92. Helsinki: SKS.

Osinsky, Pavel. 2008. War, state collapse, redistribution: Russian and German revolutions revisited. *Political Power and Social Theory* 19: 3–38.

Osinsky, Pavel, and Jari Eloranta. 2014. Why did the communists win or lose? A comparative analysis of the revolutionary civil wars in Russia, Finland, Spain, and China. *Sociological Forum* 29(2): 318–341.

Osipova, Tat'yana. 2001. *Rossiiskoe krest'ianstvo v revolyutzii i grazhdanskoi voine. [Russian peasantry in the revolution and civil war]*. Moskva: Streletz.

Pavlyuchenkov, Sergei. 1996. *Krest'yanskii Brest, ili predystoriya Bol'shevistskogo NEPa.* [The peasants' Brest, or a prehistory of the Bolsheviks' NEP]. Moskva: Russkoe kn.-vo

Payne, Stanley. 2011. *Civil war in Europe, 1905–1949*. Cambridge: Cambridge University Press.

Pipes, Richard. 1990. *The Russian revolution*. New York: Vintage Books.

Pipes, Richard. 1993. *Russia under the Bolshevik Regime*. New York: A.A. Knopf.

Pipes, Richard. 1995. *A concise history of the Russian revolution*. New York: Knopf.

Porter, Bruce. 1994. *War and the rise of the state: The military foundations of modern politics*. New York: The Free Press.

Raleigh, Donald. 2002. *Experiencing Russia's civil war: Politics, society, and revolutionary culture in Saratov, 1917–1922*. Princeton: Princeton University Press.

Rasila, Viljo. 1969. Finnish civil war and land lease problem. *The Scandinavian Economic History Review* 17: 115–135.

Retish, Aaron B. 2008. *Russia's peasants in revolution and civil war*. Cambridge: Cambridge University Press.

Riga, Liliana. 2012. *The Bolsheviks and the Russian empire*. Cambridge: Cambridge University Press.

Sanborn, Joshua. 2002. *Drafting the Russian nation: Military conscription, total war, and mass politics, 1905–1925*. DeKalb: Northern Illinois University.

Siltala, Juha. 2014. Being absorbed into an unintended war. In *The Finnish Civil War: History, memory, legacy*, ed. T. Tepora and A. Roselius, 51–89. Leiden: Brill.

Skocpol, Theda. 1979. *States and social revolutions*. Cambridge: Cambridge University Press.

Slater, Dan, and Daniel Ziblatt. 2013. The enduring indispensability of the controlled comparison. *Comparative Political Studies* 46(10): 1301–1327.

Smith, S.A. 2011. 'Moral economy' and peasant revolution in Russia: 1861–1918. *Revolutionary Russia* 24(2): 143–171.

Suny, Ronald Grigor. 1994. Revision and retreat in the historiography of 1917: Social history and its critics. *Russian Review* 53: 165–182.

Suny, Ronald Grigor. 2010. *The Soviet experiment: Russia, the USSR, and the successor states*. Oxford: Oxford University Press.

Tarrow, Sidney. 2010. The strategy of paired comparison: Toward a theory of practice. *Comparative Political Studies* 43: 230–259.

Tepora, Tuomas, and Aapo Roselius. 2014. Introduction: The Finnish Civil War, revolution and scholarship. In *The Finnish Civil War: History, memory, legacy*, ed. T. Tepora and A. Roselius, 1–18. Leiden: Brill.

Tilly, Charles. 1975. Reflections on the history of European state-making. In *The formation of national states in Western Europe*, ed. Charles Tilly, 3–82. Princeton: Princeton University Press.

Tilly, Charles. 1990. *Coercion, capital, and European states, AD 990–1992*. Cambridge, MA: Blackwell.

Upton, Anthony. 1980. *The Finnish revolution, 1917–1918*. Minneapolis: University of Minnesota Press.

Wade, Rex. 2005. *The Russian revolution, 1917*. Cambridge: Cambridge University Press.

Ylikangas, Heikki. 1993. *Tie Tampereelle: Dokumentoitu kuvaus Tampereen antautumiseen johtaneista taisteluista Suomen sisällissodassa 1918*. Helsinki: WSOY.

Repressions and Punishment Under Stalin: Evidence from the Soviet Archives

Andrei Markevich

Abstract This chapter explores patterns of repressions and punishment under Stalin using two unique datasets extracted from the Soviet archives. First, I study the profile of the Great terror at one of Soviet industrial ministries, the chief administration of metallurgy. I find that the probability of arrest was higher for party members, high-rank officials, ethnic minorities, and employees with higher education in 1937. Second, my analysis of plan fulfilment by industrial ministries during the postwar years shows that penalties were negatively correlated with production achievements. I discuss these findings in the light of political and economic explanations of Soviet repressive policy.

Keywords Punishment • Terror • Coercion • Stalin • Russia • Soviet Union

Introduction

A widespread use of severe punishment was one of key features of the Soviet system, especially under Stalin. Between 1921 and 1953, the regime executed more than a million people, imprisoned up to 20 million, exiled up to another six million, and sentenced more than 20 million to correctional works.[1] On top of that, the

[1]Figures on scales of repressions under Stalin are still imprecise. Historians continue their attempts to refine these numbers. On top of reconstruction issues, estimations depend on the definition of a political victim (Okhotin and Roginskii 2007). Stalin's political victims in the narrow sense include only those, who were repressed by the Soviet secret police. According to the document prepared in late 1953 by the secret police itself, there were about five million such persons; 800 of them were executed (Mironenko and Vert 2003, pp. 608–609). In addition, regular Soviet courts imprisoned about 14.5 million people after 1940 for political, economic, or criminal reasons (250,000 of them were executed) and sentenced about 19.5 million people to correctional works or imposed fines on them (Mironenko and Vert 2003, p. 610). The pre-1940 statistics about sentences imposed by regular courts is less precise. Figures on the flow of prisoners in labor camps suggests that up to 20 million people became Gulag prisoners between 1929, when the Soviet government started to use

A. Markevich (✉)
New Economic School, 100a Novaya Street, Skolkovo, Moscow 143025, Russia
e-mail: amarkevich@nes.ru

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*,
Studies in Economic History, DOI 10.1007/978-981-10-1605-9_5

Soviet authorities extensively used various forms of administrative and monetary punishment. The economic and historical literatures suggest two approaches to explain this policy stressing either the dictatorial nature of the Soviet regime or its reliance on the command economy in the sphere of production, i.e., highlighting either political or economic reasons of terror and punishment (e.g., see Conquest 1990; Gregory and Stuart 1999; Gregory and Lazarev 2003; Harris 2013). Despite the extensive literature, there are few studies exploring the phenomena quantitatively because of limited data accessibility.[2] This chapter is one of the first such attempts. The chapter analyzes patterns of repressions and punishment under Stalin using individual and ministry-level archival records.

I take advantage of two unique datasets extracted from the former Soviet archives and investigate two cases of implementation of Stalin's repressive policy in practice. First, I study patterns of repressions at one of Soviet industrial units, the chief administration of metallurgy (GUMP), in 1937 when Stalin's terror reached its maximum. I find that the probability of arrest at the Soviet apparatus during the Great terror depended on social and professional background of the employees. Party members, high-rank officials, ethnic minorities and employees with higher education faced higher risks. In contrast, party members with higher education – presumably a new generation of the Communist elite – had more chances to survive. Some of these characteristics might be considered as proxies for political loyalty in the eyes of the dictator. Second, I analyze relatively mild penalties imposed by the Ministry of State Control on employees at industrial ministries during the postwar years. I find that these penalties were negatively correlated with production achievements of the corresponding industrial ministries. Taken together these findings provide some support to both political and economic hypotheses of the use of repressions and punishment under Stalin.

The chapter proceeds as follows. I start with a review of the related literature on punishment and terror under Stalin. The second section analyzes the Great terror at the Chief Administration of Metallurgy. In the third section, I study penalties and plan fulfilment at industrial ministries after the Second World War. The final section concludes.

Related Literature

One explanation of the extensive use of coercion in the USSR appeals to the nature and basic characteristics of the Soviet Communist state. Indeed, from the first days in power the Bolsheviks viewed political terror as an important policy tool

forced labor in large scales, and 1953, the year of Stalin's death. About six million people were exiled, including two million former kulaks during collectivization (Polyan 2001).

[2]The only exceptions are regional case studies that compare profiles of those who were repressed in particular regions during the Great terror years with the profiles of the entire population in corresponding regions known from the 1939 population census (Ilič 2000, 2006).

against their enemies; already the first Soviet constitution adopted in 1918 stated this formally, officially defining the regime as the dictatorship of the proletariat. Under this interpretation, Stalin just extended the scope and the focus of repressions to which the regime was already predefined, while historians disagree why Stalin did that. According to the view that goes back to Leon Trotskii, one of Stalin's main political rivals, Stalin relied on terror to consolidate power and to construct his personal dictatorship, to eliminate the opposition within the party and the Soviet apparatus, replacing the so-called Lenin's guard or Old Bolsheviks, i.e., those who joined the party before the revolution and during the civil war, by a new generation of loyal Communists (e.g., see Conquest 1990). Another interpretation portrays Stalin's terror, and in particular the Great terror of 1937–1938, as a response to growing difficulties and social disorder within the country, which in their turn were caused by the implementation of the policy of rapid industrialization and forced collectivization (Getty and Naumov 1999).

Recent research in the previously closed Soviet archives has improved our knowledge on the history of Stalin's repressions. It became clear that the party was not the only or the main group targeted by the terror; the repressions affected all groups of the Soviet society. In particular, the Great terror was launched as a campaign against the Soviet elite, and from July 1937 onwards was transformed into "mass operations" against ordinary people. Stalin guided these purges personally. "Mass operations" were started and finished by special orders issued by the Soviet secret police and approved by Stalin. These orders specifically listed former kulaks, former criminals, counter-revolutionaries, and specific ethnic minorities, whose loyalty was suspect from the point of view of the dictator, as groups which had to be purged (Junge and Binner 2003; Junge et al. 2008).

Despite our greater knowledge, Stalin's repressions and the Great terror in particular still generate many controversial interpretations (for the most recent debate see chapters written by various authors in Ilič 2006 and Harris 2013). For example, historians express completely opposite views on such questions as how carefully Stalin planned and prepared the Great terror ("conducted with clear premeditation" vs. "started abruptly, without preparation" – chapters by Rees and Rittersporn in Harris 2013), how random the terror was in practice ("repressions based on categories of identities" vs. "blind" random terror – Shearer and Rittersporn in Harris 2013), and so on. There is a similar debate about reasons of variation in the scale of repressions over the years; indeed, numbers of arrests spiked in particular years, being relatively low in others. The dominant explanation of timing of the Great terror stresses Stalin's belief in a coming big war and his intention to eliminate the "fifth column" within the country and to fear potential opposition (Khlevnuk 1996; Harrison 2008; Gregory 2009). If Stalin followed this aim, even random terror against innocents could be strategically rational (Gregory et al. 2011).

An economic explanation of the wide use of repressions and punishment in the USSR points to institutional features of the command economy that tends to criminalize economic behavior because of the limited scope of other (monetary) incentives (Gregory and Stuart 2003; Belova and Gregory 2009). Indeed, the Soviet government criminalized many forms of economic behavior; the worst examples of such policy include penalties because of poor-quality production for managers (1932 decree, repeated in 1939), prison conviction because of poor labor discipline (1939 and 1940 decrees), and others. Miller and Smith (2015) argue that the whole Gulag system could be treated as an incentive device in the absence of the threat of unemployment. Under some assumptions even random penalties of innocents could represent a rational decision to deter shirking by economic agents (Harrison 2002; Lskavyan 2007). It is more difficult to explain time changes in Stalin's repressions with economic approach. In particular, a careful study of the Soviet economy before the Great terror did not reveal any deterioration in plan fulfilment or in Soviet economic development (Davis 2006).

Repressions at the Chief Administration of Metallurgy in 1937

Who became victims of Stalin's terror and who did not? Did the victims and the survivors differ from each other and if "yes", how? A systematic comparison of the two groups would help to understand better the implementation of Stalin's repression in practice; in particular, it would allow addressing statistically the widely debated issue of the randomness of repressions under Stalin.

In this section I investigate repressions at the central apparatus of the Chief Administration of Metallurgy (*Glavnoe uprvalenie metallurgicheskoj promishlennosti* – GUMP) in 1937. To the best of my knowledge, this is the first case study that analyzes individual profiles of all employees at a Soviet institution during the Great terror and investigates associations between personal characteristics and the probability of arrest. A peculiarity of information available in the former Soviet archives is a potential explanation of the lack of such studies. The Soviet secret police archives provide information about the victims but not about the survivors; on the other hand, the archives of state apparatus organizations have information about their employees but normally without specific notes on repressions. It is difficult to match these two types of information and to construct a systematic dataset on both the treatment and the control groups of Stalin's repression policy. This analysis becomes possible by a unique document in the archive of the Chief Administration of Metallurgy (RGAE 4086/6/170: 55–7[3]), which lists all employees at the central apparatus of this administration who were arrested as of mid-December 1937.[4]

[3]I employ a standard system of references to Soviet archival documents below. The abbreviation stands for the name of the archive where the document could be found (full names see in the reference list to this paper), followed by the collection (*fond)* number, then the inventory (*opis)* number and the file (*delo)* number. Figures after the colon refer to the pages in the file.

[4]The list includes 33 employees with dates of their dismissal from GUMP. The dates vary from May to December 1937. The document has no title, author, or date when it was created. There is a "not for public" stamp on the document. One-third of the people from the list appear in the public *Memorial* database on victims of the Great terror (http://lists.memo.ru/), and the dates of their dismissal known from the list fit perfectly with the dates of arrests known from the database.

I match these data on repressions with a dataset of employees at the central apparatus of the chief administration in 1937 constructed from archival documents (RGAE 4086/6/172: 23–39). The employees' dataset includes information on their individual characteristics such as gender, age, ethnicity, occupation, education, social origin, and Communist party membership.

The case of the Chief Administration of Metallurgy in 1937 provides an opportunity to analyze Stalin's repressions in one of the leading Soviet industrial units at the peak of oppression policy. The central apparatus of GUMP represented an administrative unit in Moscow that ran the whole Soviet industry of metallurgy with several dozen plants and tens of thousands of employees. The apparatus itself employed only about 400 officials (not counting the department of sales, which represented a separate administrative unit with about 150 employees). This branch of industry attracted the attention of top Soviet officials, including Stalin, because of the strategic importance of pig iron and steel for the Soviet industrialization project. Because of high level of priority of GUMP, one could expect that the Great purge at this administration provides a clear example of repressive preferences of the Soviet dictator, especially at the first stage of the campaign when "the main blow of repression was directed against members of the party, mainly those who had participated in opposition parties and movements or who had shown some kind of dissent with Stalinist policies" (Khlevnuk 2006). In other words, the GUMP case provides an opportunity to explore one of political explanations of the Great terror as a policy tool to construct a new Soviet elite that would be loyal to the dictator. An analysis of repressions at GUMP also allows investigating whether particular social characteristics of an employee, which were claimed to be proxies for political disloyalty at the start of the mass operation in August 1937, affected the implementation of terror in practice.

Table 1 presents a profile of the Chief Administration of Metallurgy in 1937. The Great terror directly affected 33 out of 400 GUMP officials, i.e., 6 % of employees were arrested. This is ten times greater than on average for the USSR: about a million arrests in 1937 in the country of 170 million, i.e., about 0.6 %. The peak of arrests at GUMP happened between August and November 1937, a period that roughly corresponds to the peak of mass operations in the country.

As a body of governance, GUMP employed many party and Young Communist League (*Komsomol*) members. Their total share was seventeen percent, almost twenty times larger than the USSR share (about 1.5 million in 1937, i.e., 0.88 %). There were few GUMP officials with "former classes" origins, i.e., those who belonged to or whose parents belonged to "former bourgeois" classes – gentry and merchants. Their share did not exceed 3 %. According to the official Soviet

---

Unfortunately, the *Memorial* database is incomplete. It includes only 2.65 million victims, and this explains the lack of information on the other two-thirds of the individuals from the GUMP list in the database. The GUMP list is complete; there are no GUMP employees not from the list whose names appear in the *Memorial* dataset.

Table 1 The Chief Administration of Metallurgy in 1937: summary statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Repressed (yes=1, no=1)	399	0.06	0.23	0	1
Communist party membership	397	0.17	0.38	0	1
“Former bourgeois” classes origin	395	0.03	0.16	0	1
“Townsmen” origin	395	0.68	0.47	0	1
Worker and peasant origin	395	0.29	0.46	0	1
Higher education	399	0.45	0.5	0	1
Ethnic minorities	387	0.33	0.47	0	1
High rank	399	0.23	0.42	0	1
Age	396	37.54	9.64	18	69
Gender (male=0; female=1)	399	0.40	0.49	0	1

ideology, social origin was an important predictor of potential political disloyalty. The share of employees with workers’ or peasants’ origins was only about 30 %. The bulk of GUMP officials were of “townsmen” (*meshchanin)* origin, belonging to various urban social strati (68 %). These townsmen had better access to education before the revolution, and education was an important requirement to work at an institution running a branch of industry. Indeed, almost a half of GUMP employees had at least some years spent in college. Interestingly, ethnic minorities occupied a non-proportionally high share among the GUMP employees, about one-third. Some 90 % of them were Jews, whose higher education and active participation in the revolutionary movement offer an explanation. Males represented 60 % of employees with an average age of 38. A quarter of occupations might be classified as high-rank positions (the head of the chief administration, his deputies, heads of departments, and their deputies).

I employ a non-linear probit regression model to analyze whether patterns of repressions at the Chief Administration of Metallurgy depended on characteristics of employees. To be precise I use the following specification:

$$\begin{aligned} Prob\,(Y_i = 1) &= \Phi\,(y_i{}^* > 0) \\ y_i{}^* &= \alpha + X_i\beta + \varepsilon_i \end{aligned} \tag{1}$$

where the latent variable, y^*, is determined by X, which includes available characteristics of employees, such as dummies for party membership, “former” social origin, ethnic minority, higher education, and high rank in the Soviet hierarchy. I first analyze the effect of these variables on the probability of arrest separately and then take them all together in order to address potential omitted variables problem. In all specifications, I also control for age and gender of employees.

Table 2 presents the results. The estimated coefficients suggest that the probability of arrest were positively associated with party membership, ethnic minority status, higher education, and high rank in the Soviet hierarchy, but was not connected with individual origins. The corresponding coefficients on variables

Table 2 Repressions at the Chief Administration of Metallurgy in 1937

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Repressed					
Communist party membership	0.51**					0.47*
	[0.254]					[0.284]
High position in the Soviet hierarchy		0.81***				0.64**
		[0.248]				[0.264]
Higher education			0.77***			0.65**
			[0.260]			[0.284]
Ethnic minority				0.46**		0.43*
				[0.213]		[0.245]
Townsmen or "former bourgeois" classes origin					0.15	–0.07
					[0.247]	[0.297]
Age	0.02	0.007	0.003	0.01	0.008	0.01
	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]	[0.016]
Gender	–0.25	0.02	–0.08	–0.29	–0.33	0.27
	[0.263]	[0.294]	[0.280]	[0.260]	[0.258]	[0.329]
Number of observations	397	399	396	387	394	379

Probit regression model is employed. Independent variables, except age, are dummies. Standard errors are in brackets

**indicates p-value < 0.01, **p-value < 0.05, *p-value < 0.1

of interest do not differ much whether I take these variables into regressions separately or at once. All variables on which coefficients are statistically significant in columns 1–5 remain statistically significant in column 6 at least at the 5 % level; their magnitudes remain similar as well. The coefficients on the dummy for social origins are always statistically insignificant (columns 2 and 6). The marginal effects estimated for the specification reported in the last column of Table 2, i.e., controlling on all characteristics for which data are available, were substantial for an average employee. Party membership increased the probability of repressions by 4.6 percentage points, or by 84 % (an average probability of arrest was 5.5 %; 84 = 4.6/5.5). Similarly, high-ranking officials were in the risk group. Their occupation increased the probability of arrest by 6.6 percentage points or by 120 % (=6.6/5.5). Higher education increased probability of victimhood by 5.3 percentage points or by 96 % (=5.3/5.5). Finally, arrests were also 3.7 percentage points more likely for ethnic minorities or by 67 % (=3.7/5.5).

These results provide support for the interpretation of the Great terror as a campaign conducted by the dictator to replace the existing elite with a new one. Indeed, both the party membership and a high rank in the Soviet hierarchy doubled the probability of repression for an average GUMP employee. Higher education as a predictor of arrest might be viewed as further evidence in support of this interpretation. A coalition of party leaders together with non-party "old" technical specialists inherited by the Communists from the previous regime ran Soviet industry in the mid-1930s. For years after the revolution, the Bolsheviks had to rely on "old" specialists because of a shortage of their own cadres with higher education. The Great terror changed this policy. I explore political motives of Stalin's oppression policy further in Table 3.

Patterns of repressions at GUMP also highlight ethnic minorities as another group of risk that could be interpreted as additional evidence on political origins of the Great terror. Indeed, the Great purges had an ethnic component, so called mass operations against specific nationalities whose loyalty was considered problematic because of cross-border ties with "mother" states (Germany for the Soviet Germans, Poland for the Soviet Poles, and so on). Interestingly, however, in the case of GUMP repressions, ethnical minorities suffering from the terror were not those, which were listed in the 1937 mass operation decrees. Ethnical minority victims in GUMP were mainly Jews who were not targeted by national campaigns, at least officially. With the data currently available, it is unclear whether the Great terror had other ethnic dimensions on top of national campaigns and specifically affected Jews, widely represented in the Soviet state apparatus after the 1917 revolution. Nevertheless, it is clear that the case of GUMP was not the only example of such policy. The Great terror had similar consequences for the apparatus of the secret police itself, where the share of Jews decreased from more than a third in 1937 to less than 4 % in 1940 (Petrov and Skorkin 1999).

In Table 3 I analyze the hypotheses that the Great terror specifically targeted the GUMP elite, i.e., the old Bolsheviks and old specialists, clearing positions for a new generation of the Communist leaders who at the same time had sufficient education to run Soviet industry. Old Bolsheviks represented those party members who joined the party at least before 1920. Unfortunately, I do not have a good proxy for this group since I do not have data on the number of years of particular individuals in the party. In its place, I employ an interaction of the party membership and high-rank position dummy since the bulk of old Bolsheviks occupied high ranks in the Soviet hierarchy after the revolution. I add this interaction into the specification reported in column 6 of Table 2 and present the results in column 1 of Table 3. The interaction coefficient is positive but is not statistically different from zero, a possible consequence of the quality of my proxy. In column 2 of Table 3 I run the same regression but with the interaction of the party membership with higher education dummy as a proxy for the new generation of the Soviet elite. The resulting coefficient is negative and statistically significant at the 10 % level. If an average employee belonged to this group, her probability of arrest was 3 percentage points (55 %) lower. An alternative interpretation of the latter results is that the probability of arrest for old specialists, i.e., non-party employees with higher education, was higher during the Great terror.

Table 3 Party membership and repressions at the Chief Administration of Metallurgy

	(1)	(2)
Variables	Repressed	
Townsmen or "former bourgeois" classes origin	–0.07	–0.09
	[0.298]	[0.306]
Communist party membership	0.43	1.24**
	[0.396]	[0.518]
Ethnic minority	0.43*	0.41
	[0.245]	[0.250]
Higher education	0.65**	–0.03
	[0.284]	[0.451]
High rank	0.62**	0.65**
	[0.298]	[0.268]
High rank X Party	0.08	
	[0.524]	
Higher education X Party		–1.12*
		[0.598]
Age	0.01	0.02
	[0.016]	[0.017]
Gender	0.28	0.35
	[0.329]	[0.340]
Number of observations	379	379

Probit regression model is employed. Standard errors are in brackets. Independent variables, except age, are dummies
**p-value < 0.05, *p-value < 0.1

Plan Fulfilment and Penalties in the Soviet Union After the Second World War

In this section I alter the angle of my analysis as well as the dataset to explore the economic foundations of punishment under Stalin. Past empirical research on this question has been hampered by the lack of data on economic achievements and punishment in the Soviet system. I managed to overcome this problem by combining data on annual plan fulfilment by Soviet industrial ministries in 1946–1952 from the leading Soviet newspapers *Pravda* and *Izvestia* (partially summarized in Zaleski 1980)[5] with archival information on penalties imposed on officials employed at these ministries by the Ministry of State Control (GARF 8300/2/134; 317; 615; 767).

According to the economic hypothesis, punishment was an incentive to motivate Soviet managers to work hard and to fulfil economic plans assigned by the government. Indeed, the Soviet authorities declared personal responsibility of Soviet

[5] I thank Mark Harrison for bringing my attention to these data.

officials for the achievements of the units of which they were in charge. Stalin's closest economic subordinate, Sergo Ordzhonikidze, the minister of heavy industry in the 1930s, formulated this explicitly in a public speech in the first half of the decade: "That director, that engineer, that shop boss, that technical director who does not insure that the directives of the commissariat, the government, or the central committee of our party [including plan directives] are not fulfilled exactly is not a director" (RGAE 7297/38/106: 12; June, 16 1934). The plan was the law, and the violators of the law were supposed to be punished (Nove 1958; Belova and Gregory 2002). However, evidence of whether this approach was systematically realized in practice are mixed (Gregory 1990). For example, according to the Soviet rules, production of poor-quality goods was a crime, but this rule was rarely implemented in practice (Solomon 1998) despite multiple cases known to the authorities (Markevich and Harrison 2006). The constructed dataset on plan-fulfilment by industrial ministries allows addressing the question on the use of punishment as an incentive systematically.

There were a number of institutions in the Soviet Union whose task was to monitor efforts and achievements of state industrial units. The Ministry of State Control together with the party control organs, special units of the Ministry of Internal Affairs, and the prosecution service were the key bodies in the Soviet control system. They had the right to penalize economic agents for poor results. The government also empowered these organizations with the privilege of intervening in production processes of the economic agents they monitored. They could assist agents to implement production plans if necessary (Markevich 2011). For example, according to the 1951 statute of the Ministry of State Control, its primary task was to monitor economic activities of industrial units, to penalize wrongdoers with criminal, administrative sentences or fines and to intervene to secure realization of main production tasks. If the ministry believed that a massive intervention, which would substantially change original plans, was required, the ministry could appeal to the government, and the government could itself conduct such actions (GARF 8300/2/1: 104–115).

The activities of the control organs were organized along the hierarchical lines of the Soviet state apparatus. In particular in the Ministry of State Control, departments were responsible for particular branches administered by industrial ministries, i.e., each department monitored and inspected one industrial ministry. Annual reports of the Ministry of State Control for 1946–1952 provide information on activities of each department – number of imposed penalties, number of interventions, and so on.

Table 4 reports summary statistics of the constructed dataset where an industrial ministry in a particular year is a unit of observation. The panel is unbalanced because of permanent reorganizations of the Soviet industrial hierarchy (Crowfoot and Harrison 1990). The panel includes 144 observations on 41 ministries during eight years. I deliberately excluded the outliers of extremely low and high plan fulfilment from further consideration, namely below 90 % and above 110 %, because they do not tell much about achievements of the corresponding industrial units rather than signal that something wrong happened at a stage of elaboration of plans.

Table 4 Plan fulfilment and penalties imposed on Soviet industrial ministries by the Ministry of State Control in 1946–1952: summary statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Plan fulfilment (percent)	144	102.07	3.5	92	109
Penalties of all types	144	38.31	34.37	0	226
Criminal sentences	144	2.62	4.61	0	28
Administrative penalties	144	24.01	21.14	0	137
Fines	144	11.67	12.83	0	74
Ministry of State Control orders	144	12.6	11.87	0	98
Ministry of State Control interventions	120	12.52	15.04	0	79
Government interventions	120	8.23	7.47	0	44
Number of production units in a corresponding industrial ministry checked by the Ministry of State Control	71	57.41	28.05	7	141

After the war, an average ministry over-fulfilled its annual plan by 2 %, and about 38 ministerial officials received penalties from the Ministry of State Control during an average year. Twenty-four out of these 38 penalties were administrative (reprimands, "strong" reprimands, and so on); they affected career prospects of the employees. Twelve were fines and three represented criminal sentences. There was substantial variation in penalties across the ministries and over the years. There were ministries with zero penalties in particular years, and the maximum was more than 200 penalties.

The number of orders issued by the Ministry of State Control related to activities of industrial ministries reflects the level of attention of the control organs to each ministry (GAFR 8300/2/1: 108), i.e., represent a proxy for resources spent on monitoring. The Ministry of State Control issued on average one order per month for each industrial ministry. The number of industrial units (factories, constructions, research institutions, and similar) subordinated to a corresponding industrial ministry and inspected by the Ministry of State Control in a particular year is another proxy for the scope of monitoring. The Ministry of State Control inspected about 70 industrial units in an average industrial ministry in an average year. Unfortunately, these figures are known only for four years (1946, 1947, 1948, and 1950).

As mentioned above, the Ministry of State Control could also intervene in production processes adjusting original production plans of industrial ministries. There were on average about twelve direct interventions per ministry in a year. On top of that, the Ministry of State Control initiated about eight governmental interventions that affected an average ministry in a year.

I use the constructed dataset to explore the relationship between penalties and plan fulfillment. I employ the level of punishment as a dependent variable and the level of plan fulfilment as an independent variable in a linear model controlling for ministry and year fixed effects as well as other available controls. In particular, I estimate the following model:

$$Penalties_{it} = \alpha + \beta^{*} PlanFulfillment_{it} + \gamma^{*} \left(Controls_{it}\right) + \left(Ministry_{i}\right) + \left(Year_{t}\right) + \varepsilon_{it} \quad (2)$$

where subscripts i and t index industrial ministries and years, respectively. *Penalties* are either number of all penalties or penalties by type imposed by the Ministry of State Control on officials from a particular industrial ministry. *PlanFulfillment* is the main variable of interest that reveals a percentage of a plan fulfilled by a particular ministry in a particular year. *Controls* stand for a vector of controls reflecting the scope and the scale of monitoring activities by the Ministry of State Control and its direct and indirect interventions. With *Ministry* and *Year* variables I incorporate ministry and year dummies to account for potential cross-sectional heterogeneity and time shocks. I cluster standard errors by ministry.

The estimating model suffers from a potential endogeneity problem because of possible reverse causality and omitted variables. Indeed, extensive punishment could decrease plan fulfillment, as in the interpretation of the Great terror by Davis (2006). I also do not have information on types of offenses being penalized to control for in the regression analysis. This might be important because Soviet managers often preferred to violate some minor rules in a very complicated Soviet set of laws regulating economic activities in order to secure implementation of the main goal of gross industrial output, a strategy known as the "honest manager dilemma" (Belova and Gregory 2002).[6] So I cannot draw conclusions about causality in the explored relationship; nevertheless, correlations are already informative and interesting.

Table 5 presents the results. In the first column I regress penalties on plan fulfilment, ministry, and year fixed effects. The coefficient on the variable of interest is negative but imprecisely estimated. It becomes statistically significant once I control for monitoring resources, adding number of orders issued by the Ministry of State Control related to the activities of the corresponding industrial ministry (column 2). The magnitude of the effect is the following: an increase of plan fulfilment by 1 percentage point was associated with a decrease in punishment by 0.6 penalties. The negative correlation between plan fulfilment and punishment survives if I control for direct and indirect interventions of the Ministry of State Control into activities of the industrial ministries it monitored (column 3). The correlation is also stable if I add another proxy of the scope of monitoring, namely the number of production units checked by the Ministry of State Control (column 4).

In Table 6 I repeat the exercises from Table 5 but look at penalties by type as the dependent variable. I report the first and the last specifications from Table 5 for each type. The negative correlation between punishment and plan fulfilment is driven mainly by administrative penalties (columns 3 and 4) rather than criminal sentences (columns 1 and 2). The correlation between fines and plan fulfilment

[6] Petr Parshin, the Soviet minister of machine-building and instrument-making, noticed at a meeting in the Ministry of State Control in 1946, "If you [inspectors] inspect formally, you will be able to reprimand any of us" (GARF 8300/2/189: 55).

Table 5 Correlates of penalties imposed on Soviet industrial ministries in 1946–1952

	(1)	(2)	(3)	(4)
Variables	All penalties			
Plan fulfilment	–0.71	–0.59*	–0.61**	–1.1***
	[0.716]	[0.303]	[0.256]	[0.380]
Ministry of State Control orders		2.13***	2.36***	3.03***
		[0.203]	[0.203]	[0.240]
Government interventions			1.16**	0.34
			[0.472]	[0.486]
Ministry of State Control interventions			–0.09	–0.18*
			[0.055]	[0.093]
Number of production units checked				0.04
				[0.034]
Year fixed effects	Yes	Yes	Yes	Yes
Ministry fixed effects	Yes	Yes	Yes	Yes
Observations	144	144	120	71

OLS regression model is employed. Standard errors are clustered by ministry and are reported in brackets. *Plan fulfillment* is a percent of implementation of plan output task by an industrial ministry. *Ministry of State Control orders* is a number of orders issued by the Ministry of State control in respect to activities of a corresponding industrial ministry. *Government* and *Ministry of State Control interventions* are number of interventions undertook by these bodies in respect to a corresponding industrial ministry. *Number of production units checked* is a number of production units in a corresponding industrial ministry whose activities were checked by the Ministry of State Control during a corresponding year

***indicates p-value < 0.01, **p-value < 0.05, *p-value < 0.1

is also negative and becomes significant once I include additional controls into specification (columns 5 and 6).

The obtained results do not contradict the hypothesis on economic origins of punishment in the Soviet Union, while they do not allow making a strong conclusion about the causal relationship between poor economic achievements and penalties. However, if there was such relationship, these results show that the penalties for economic misbehavior was relatively light; the control organs used administrative reprimands or fines rather than criminal sentences.

Conclusion

The two cases studies explored above do not allow distinguishing between the economic and political explanations of the causes of repressions and punishment under Stalin. They provide some evidence in support of both. The case study on repressions at GUMP shows that social and ethnic groups whose loyalty was questionable in the eyes of the dictator saw higher risks of arrests. The negative correlation between plan fulfilment and penalties in industrial ministries after

Table 6 Penalties by type and plan fulfilment

	(1)	(2)	(3)	(4)	(5)	(6)
	Criminal sentences		Administrative penalties		Fines	
Plan fulfillment	0.04	0.1	–0.63***	–0.58**	–0.008	–0.61*
	[0.110]	[0.235]	[0.221]	[0.213]	[0.206]	[0.322]
Ministry of State Control orders	0.34***	0.39**	1.08***	1.87***	0.83***	1.06***
	[0.054]	[0.150]	[0.183]	[0.116]	[0.168]	[0.222]
Government interventions		0.03		0.53**		–0.22
		[0.189]		[0.260]		[0.364]
Ministry of State Control interventions		–0.08		–0.07		–0.07
		[0.060]		[0.061]		[0.080]
Number of investigated agents		–0.004		–0.02		0.06**
		[0.013]		[0.021]		[0.028]
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Ministry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	144	71	144	71	144	71

OLS regression model is employed. Standard errors are clustered by ministry and are reported in brackets. *Plan fulfillment* is a percent of implementation of plan output task by an industrial ministry. *Ministry of State Control orders* is a number of orders issued by the Ministry of State control in respect to activities of a corresponding industrial ministry. *Government* and *Ministry of State Control interventions* are number of interventions undertook by these bodies in respect to a corresponding industrial ministry. *Number of production units checked* is a number of production units in a corresponding industrial ministry whose activities were checked by the Ministry of State Control during a corresponding year

***indicates p-value < 0.01, **p-value < 0.05, *p-value < 0.1

the war does not contradict the economic hypothesis of the origins of Soviet punishment. I cannot make a stronger conclusion on the relative importance of the two hypotheses; that would require better data and further research.

With my findings on patterns of repressions at the Chief Administration of the Metallurgy in 1937, I contribute to the empirical literature on the Great terror (Vatlin 2004; Junge et al. 2008). This is the first empirical study on repressions at a Soviet central apparatus institution. The profile of repression at GUMP supports the interpretation of the Great terror as a campaign against the existing elite, who ran Soviet industry in the mid-1930s, and its replacement by a new generation of Communists loyal to Stalin. I also find that the terror affected all ethnic minorities, and was not limited to minorities listed in the secret police orders on so-called national campaigns. GUMP was one of the leading institutions in the Soviet industrial hierarchy; however, it is less clear whether the GUMP patterns could be extrapolated to the whole country.

I also contribute to the literature on incentives in the command system. I provide evidence that does not contradict to the hypothesis that punishment followed poor plan fulfillment. It is less clear how efficient such policy was. In particular, I found that punishment, which was correlated with poor plan fulfillment, was relatively light and affected career prospects of Soviet managers rather than produced a negative effect on their current utility. In addition, officials could spend their efforts to revise plans instead of to fulfil them; this would allow them to escape punishment without increasing output (Berliner 1957; Zaleski 1980; Gregory 1990). The question on the efficiency of punishment as an incentive in the Soviet command economy requires further research.

Acknowledgments I would like to thank the Hoover Institution, where I was a National Fellow in 2014/2015 academic year, for its hospitality and support. I am also grateful to Paul Gregory, Mark Harrison, and Sergei Guriev for their comments. A standard disclaimer is applied.

Archival and Published Primary Sources

GARF – State archive of Russian Federation (Moscow).
RGAE – Russian State archive of national economy (Moscow).

References

Belova, Eugenia, and Paul R. Gregory. 2002. Dictator, loyal, and opportunistic agents. *Public Choice* 113(3–4): 265–286.

Belova, Eugenia, and Paul R. Gregory. 2009. Political economy of crime and punishment under Stalin. *Public Choice* 140: 463–478.

Berliner, Joseph S. 1957. *Factory and manager in the USSR*. Cambridge, MA: Harvard University Press.

Conquest, Robert. 1990. *The great terror: A reassessment*. Oxford: Oxford University Press.

Crowfoot, John, and Mark Harrison. 1990. The USSR Council of Ministers under late Stalinism, 1945–54: Its production branch composition and the requirements of national economy and policy. *Soviet Studies* 42(1): 41–60.

Davis, Robert. 2006. The Soviet economy and the launching of the Great Terror. In *Stalin's terror revisited*, ed. Melanie Ilič. Houndmills: Palgrave Macmillan.

Getty, Arch, and Oleg Naumov. 1999. *The road to terror. Stalin and the self-destruction of the Bolsheviks, 1932–1939*. New Haven: Yale University Press.

Gregory, Paul R. 1990. *Restructuring the Soviet economic bureaucracy. Soviet interview project series*. Cambridge: Cambridge University Press.

Gregory, Paul, and Valery Lazarev (eds.). 2003. *The economics of forced labor: The Soviet Gulag*. Stanford: Hoover Institution Press.

Gregory, Paul R. 2009. *Terror by quota. State security from Lenin to Stalin an archival study*. New Haven: Yale University Press.

Gregory, Paul, and Robert Stuart. 1999. *Comparative economic systems*. Boston: Houghton Mifflin.

Gregory, Paul, and Stuart Robert. 2003. *Comparative economic systems*, 7th ed. Boston: Houghton Mifflin.

Gregory, Paul, Phillip Schroder, and Konstantin Sonin. 2011. Rational dictator and the killing innocent: Data from Stalin's archives. *Journal of Comparative Economics* 39: 34–42.

Harris, James (ed.). 2013. *The anatomy of terror. Political violence under Stalin*. Oxford: Oxford University Press.

Harrison, Mark. 2002. Coercion, compliance, and the collapse of the Soviet command economy. *Economic History Review* 55(3): 397–433.
Harrison, Mark (ed.). 2008. *Guns and rubles. The defense industry in the Stalinist State*. New Haven: Yale University Press.
Ilič, Melanie. 2000. The Great Terror in Leningrad: A quantitative analysis. *Europe-Asia Studies* 52(8): 1515–1534.
Ilič, Melanie (ed.). 2006. *Stalin's terror revisited*. Houndmills: Palgrave Macmillan.
Junge, Mark, and Rolf Binner. 2003. *Kak terror stal bolshim. Sekretnij prikaz N004447 i telhnologii ego ispolneniya* [How terror became great. A history of the secret odere N00447 and mechanisms of its implementation]. Moscow: AIRO.
Junge, Mark, Gennadij Bordukov, and Rolf Binner. 2008. *Vertikal' Bolshogo terrora. Istoriya operatsii po prikazu NKVD N00447* [Vertical of Great Terror. A history of the operation of NKVD order N00447]. Moscow.
Khlevnuk, Oleg. 1996. *Politburo. Mekhanizmi vlasti v 1930-e godi.* [Politburo. Mechanisms of Power in the 1930s]. Moscow: Rosspen.
Khlevnuk, Oleg. 2006. Economic officials in the Great Terror. In *Stalin's terror revisited*, ed. Melanie Ilič. Houndmills: Palgrave Macmillan.
Lskavyan, Vahe. 2007. A rational choice explanation for Stalin's 'Great Terror'. *Economics and Politics* 19(2): 259–287.
Markevich, Andrei. 2011. How much control is enough? Monitoring and interventions in the Stalinist command economy? *Europe-Asia Studies* 63(8): 1449–1468.
Markevich, Andrei, and Harrison Mark. 2006. Quality, experience, and monopoly: The Soviet market for weapons under Stalin. *The Economic History Review* 59(1): 113–142.
Miller, Marcus, and Jennifer Smith. 2015. In the shadow of the Gulag: Worker discipline under Stalin. *Journal of Comparative Economics* 43(3): 531–548.
Mironenko, S.V., and N. Vert. (eds.). 2003. *Istoriya Stalinskogo Gulaga* [A history of Stalin's gulag], Massoviae repressii v SSSR [Mass repressions in the USSR], vol. 1. Moscow: Rosspen.
Nove, Alec. 1958. The problem of "Success Indicators" in Soviet industry. *Economica,* New Series 25(97): 1–13.
Okhotin, N.G., and A.B. Roginskii. 2007. *O mashtabe politicheskikh repressii v SSSR pri Staline: 1921–1953* [About the scales of political repressions in the USSR under Stalin] http://demoscope.ru/weekly/2007/0313/analit01.php. *Mimeo.*
Petrov N.V., and K.V. Skorkin. 1999. *Kto rukovodil NKVD, 1934–1941. Spravochnik*. [Who run NKVD, 1934–1941. An index]. Moscow.
Polyan, Pavel. 2001. *Ne po svoej vole: istoriya I geografiya prinuditelnoj migratsii v SSSR* [Against their wills: History and geography of forced migrations in the USSR]. Moscow: OGI.
Solomon, Peter. 1998. *Sovetskaya ustitziya pri Staline* [The Soviet Justice under Stalin]. Moscow: Rosspen.
Vatlin, A.Yu. 2004. *Terror raionnogo mashtaba: Massovie operatsii NKVD v Kuntsevskom raione Moskovskoj oblasti 1937–1938 gg* [Terror in a district: Mass operations of NKVD in Kuntsevo district of the Moscow region in 1937–1938 gg]. Moscow: Rosspen.
Zaleski, Eugene. 1980. *Stalinist planning for economic growth, 1933–1952*. Chapel Hill: The University of North Carolina Press.

Part III
Economic History of the World Wars

Myths of the Great War

Mark Harrison

Abstract There are persistent myths about every aspect of the Great War of 1914–1918: how it began, how it was won, how it was lost, and how the peace was made. I consider and reject the arguments that the war broke out inadvertently, that the Western front saw needless slaughter, that Germany was starved out of the war by the food weapon, and that the peace treaty that ended the war was the cause of another war.

Keywords Blockade • Mobilization • Great War • Hyperinflation • Interdependence • Rational calculation • Strategic interaction • Reparations • War of attrition

Introduction

One hundred years later, perceptions of the Great War continue to resonate in today's world of international politics and policy.[1] Most obviously, does China's rise show a parallel with Germany's a century ago? Will China's rise, unlike Germany's, remain peaceful? The *Financial Times* journalist Gideon Rachman wrote:

> The analogy [of China today] with Germany before the First World War is striking ... It is, at least, encouraging that the Chinese leadership has made an intense study of the rise of

[1]This paper was presented as keynote lectures to the Economic History Society annual conference at the University of Warwick, 28 March 2014, and the Ninth Appalachian Spring Conference in World History and Economics, Appalachian State University, Boone, North Carolina, 12 April 2014; and to meetings of the Lancaster University Economics Society on 25 November 2014 and the Defence Economics Seminar at the Ministry of Defence, London, on 25 March 2015. I thank the organizers and participants for comments and questions; and Karen Brandon, Dan Bernhardt, Nick Crafts, Ian Gazeley, Vlad Kontorovich, Will O'Neill, Avner Offer, Carlo Ristuccia, and Stephen Schuker for advice and discussion.

M. Harrison (✉)
Department of Economics, University of Warwick, Coventry CV4 7AL, UK
e-mail: mark.harrison@warwick.ac.uk

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_6

> great powers over the ages—and is determined to avoid the mistakes of both Germany and Japan."[2]

The idea that China's leaders wish to avoid Germany's mistakes sounds encouraging, certainly. But what are the "mistakes," exactly, that they will now seek to avoid? It is not encouraging if social scientists and historians remain uncertain what mistakes were made and even whether they were mistakes in the first place.

While attention has been focused on China's parallel with the rise of Germany, Russia's present decline in some ways resembles that of Austria–Hungary and has no less disturbing implications: a multi-national empire struggling to manage a fall from past greatness in the face of rising ethnic tensions and powerful rivals competing for influence in bordering states.

The myths of the Great War challenge the skills of both historians and economists. Historians face the challenge of preserving and extending the record and contesting its interpretation—especially when reasonable people differ over the meaning. If anything challenges the economist, it is surely persistence in behavior that is both costly and apparently futile or self-defeating.

Closer study of the historical record of the Great War reveals a story full of foresight, intention, calculation, and causation. Some consequences that are commonly thought to have been unintended were considered beforehand and fully discounted; others were not consequences at all.

Myth #1. How the War Began

Interviewed earlier this year at Davos, Japanese premier Shinzo Abe likened China and Japan today to Britain and Germany in 1913.[3] He commented on the similarity of their rivalry. He noted that they shared a strong trading relationship and that a century ago this had not prevented strategic tensions leading to the outbreak of conflict. Today, he concluded, any "inadvertent" conflict would be a disaster. A similar impression is conveyed by Christopher Clark's bestselling account of the outbreak of war in 1914, *The Sleepwalkers*.

The idea of an inadvertent war is sometimes traced to Lloyd George. In 1918 Lloyd George was in no doubt who bore responsibility for the war; he demanded "Trial of the Kaiser" (Purcell 2006: 75). By 1920 he had changed his mind. In a later memoir (Lloyd George 1938: 32) he famously wrote:

---

[2] "The shadow of 1914 falls over the Pacific," *Financial Times*, February 4, 2013.

[3] "Davos leaders: Shinzo Abe on WW1 parallels, economics and women at work," *Financial Times*, January 22, 2014. Subsequently the Chinese foreign minister Wang Yi was quoted as rejecting the parallel on the following grounds: "1. The parallel is a misleading one. Military conflict is now unthinkable. 2. Japan caused the Second World War. This is unambiguous." "Davos: China rejects Abe's WW1 analogy," *Financial Times*, January 24, 2014.

> The nations slithered over the brink into the boiling cauldron of war without any trace of apprehension or dismay.

In fact, several versions of the Great War now in circulation deny or qualify agency. These are stories of nationalism, imperialism, and coordination failure. According to them, national leaders were trapped into actions they did not intend by commercial interests, the demands of the mob, and alliance commitments.

Faced with the hypothesis of an "inadvertent" conflict, the social scientist has many questions. When the actors decided on war, did they not calculate their actions or intend the results? The economist's standard model of strategic interaction demands evidence of individual agency (rather than of unconscious collective drives), of unbiased "rational" expectations, and of backward induction of one's own best choice based on the expected best choice of the adversary.

It is a myth that such thought processes were absent from the decision for war. On the contrary the record shows that the war was brought about largely by design, and among those that designed it there was realistic foresight of the scale, scope, character, duration, and even outcome of the war.

In every country the decision for war was made by a handful—literally—of people at the apex of each political system (Hamilton and Herwig 2004: 238–241). Their councils were "saturated with agency" (Clark 2012: xxvii). The cliques themselves were not united, so that there were also waverers in every country including the German, Austrian, and Russian emperors, the German premier Bethmann Holweg, and the British finance minister (later premier) Lloyd George. At crucial moments, however, those that favored war were able to sway the others.

An implication is that the war was not inevitable. Minor variations in the course of events and the personalities involved might have had different outcomes. The Sarajevo assassination killed a voice for peace in Vienna and also made it more difficult for similar voices to be heard (Fromkin 2007: 154). But, in the powerful words of Margaret MacMillan (2013: 605): "There are always choices."

No one was swayed by commercial interests, which were against the war in all countries (Hamilton and Herwig 2004: 241–248), or by public opinion more widely, which was taken by surprise (Ferguson 1998: 174–201). Public opinion was considered, only to bolster the legitimacy of the actions the actors had decided to take anyway. In Germany, for example, Moltke (cited by MacMillan 2013: 480) wrote to Bethmann Holweg after the so-called imperial war council of December 1912 that "we can … face even the most difficult tasks with confidence, if we manage to formulate the casus belli in such a way that the nation will take up arms unitedly and enthusiastically." No one was trapped into war by alliance commitments. Instead, they considered carefully whether or not to honor them, or even went beyond them. Thus in its "blank check" to Austria, Germany went far beyond its alliance obligation. Italy, in contrast, went to war in 1915 against its former allies. Nor were they trapped into war by the pressure of mobilization timetables; in both Berlin and St. Petersburg, the war advocates exploited the timetables to force the waverers to commit to war (e.g. Herwig 1997: 25).

What ruled the calculation in every country was the national interest as they perceived it. But on what was the "national interest" based? As everywhere, on shared beliefs and values.[4] These began with national identity, in which the well-being of the nation was commonly identified with persistence of the ruling order. They extended to shared values of power, status, honor, and influence, and then to shared beliefs about the forces underlying the distribution of power in the world. Strikingly, the decision makers in every country were subscribers to a virtual world where the zero-sum game of power was being played out, not the positive-sum game of commerce and development.

There is clear evidence that some of these actors had a specific intent to bring about a war. Two things muddy the water. One is the efforts made later to destroy the evidence and distort the record (Herwig 1987). Another is that those that intended war did not have the same war in mind—although they still understood the wider conflict that could follow (Fromkin 2007). In Vienna, Chief of the General Staff Conrad and Foreign Minister Berchtold intended war with Serbia in order to assert the integrity of the Austro–Hungarian Empire—but they knew that the Russians might intervene and so widen the conflict. In Berlin, Chief of the General Staff Moltke and War Minister Falkenhayn planned war with Russia before the Russian rearmament would be completed—but they knew that this would also entail war with France.

All this provides evidence that in 1914 the great powers followed the principles of strategic interaction, including backward induction by which all players choose their own best moves, taking into account the adversary's likely response. Within the governing cliques of the Great Powers, each had reasonable understanding of the others' war plans, based on open signals and confirmed by covert intelligence (Macmillan 2013: 314–352). That is, each government shared a broad understanding that, if Austria attacked Serbia, Russia would probably mobilize against Austria and Germany; Germany would probably attack France as well as Russia; and Britain would probably come in on the side of France (Herwig 2002). Austrian foreign minister Berchtold promoted aggression "even though our operations against Serbia should bring about the great war." From Berlin, Wilhelm II told Vienna to count on "Germany's full support"; he wrote that Germany fully expected war with Russia and for years had made all preparations with this in mind.

Each country's likely reaction was not known with certainty, and there are well-known moments when they were misread. But the theory of deterrence (Schelling 1966; 92–125) does not require certainty; indeed, deterrence is thought to be more effective when each side retains some discretion. Of course, leaving the adversary in a state of uncertainty is not the same thing as being uncertain oneself, and the latter condition reflected the influence of the waverers.

To bring war about, they also encouraged each other: when the Germans encouraged the Austrians to make war on Serbia in July 1914 among them were

[4]On the "national interest" in 1914 see Hamilton and Herwig (2004: 239); on the leaders' "shared political culture" see Clark (2013: 560).

those that expected this would provide the best opportunity to attack France and Russia (Hamilton and Herwig 2004; Fromkin 2007). Similarly, the Russians and French egged each other on, although the Russians had their eyes on Austria and the French on Germany (McMeekin 2011: 54).

Is it true that everyone expected a short war? According to a *Financial Times* editorial for New Year's Day, 2014:

> In 1914 some European politicians and generals, their outlook shaped by the limited wars that had unified Germany and Italy half a century earlier, harboured this illusion.[5]

But it was not, in the main, "politicians and generals" that suffered from the illusion. It was the ordinary uninformed citizens that expected a short war. The "short war" illusion arose from the warnings of Bloch (1899) and Angell (1910) about the destructive force of modern warfare and the dependence of prosperity on economic interdependence. The military had heard the warning and had discounted it (Macmillan 2013: 305–306). In reality the idea of a short war was not so much a shared illusion as a shared hope: starting from Schlieffen, everyone hoped the war could be short.

Military planners were ready not only for a great war but also for a long one. Schlieffen's own staff warned that quick victory might well be replaced by "a tedious and bloody crawling forward step-by-step" (Clark 2013: 561). Updating German war plans in 1905 and 1906, Moltke himself envisaged one possible outcome as "a people's war, one which would not be concluded in a single battle" but a "murderous European war," a "general European massacre, at whose horror one could only shudder," a "long and protracted struggle" that would continue until "the peoples' energy had been entirely broken"; if victorious, Germany would still be "exhausted in the extreme" (Herwig 2002: 688–692). Preparing for war, German administrators planned how to feed the population under blockade (Lee 1975). Moltke himself explained that respecting Dutch neutrality would provide Germany's "wind pipe," or neutral channel to overseas trade (Herwig 2002: 689. These considerations made absolutely no sense if they planned only for a short war.

Across the Channel in August 1914, while some others thought the war might be a brief commitment, War Secretary Kitchener thought "if things go wrong the war might last two or three years at least" (cited by French 1988: 387). Prime minister Asquith anticipated "Armageddon." French and Russian generals looked forward to the "extinction of civilization" (Clark 2013: 561).

Was irrational over-optimism a factor? In this context over-optimism would have a precise meaning: that the sum of probabilities of expected victory among the great powers would exceed one. Evidence of over-optimism is strikingly absent among those that brought the war about in Berlin and Vienna. Possibly there was over-optimism in St. Petersburg: both Russians and Germans overestimated the extent to which Russia was already stronger than Germany. In the present, however, both German and Austrian leaders had clear premonitions of defeat (Berghahn

[5] "Reflections on the Great War," *Financial Times*, January 1, 2014.

1973, 2013; Ferguson 1998: 13; 2005: 19). As war began, German War Minister Falkenhayn put it: "Even if we go under as a result of this, it still was beautiful." In Vienna Conrad told his mistress: "It will be a hopeless struggle, but ... such an ancient monarchy and such an ancient army cannot perish ingloriously." The Austrian Kaiser Franz Joseph wrote: "If we must go under, we better go under decently" (from Herwig 1997: 11, 22, 37).

Far from optimism, their attitude is better described as *rational pessimism*: they did not expect victory, but they did evaluate the expected payoff from remaining at peace as worse than that from war. Thus Moltke told Conrad on 12 May 1914: "To wait any longer means a diminishing of our chances" (from Herwig 1997: 51). Bethmann Holweg to Kurt Riezler, 7 July 2014: "Russia's military power growing fast ... Austria grows ever weaker ... The future belongs to Russia, which grows and grows into an ever greater weight pressing down on our chest" (from Erdmann 1972: 181–93).

Could the Great War have been avoided? The case continues to be made that avoidance of war in such circumstances can be achieved by mediation and accommodation. Gideon Rachman, for example, has contrasted "Munich" and "Sarajevo":

> If leaders warn against "another Munich", they are almost always advocating a tough response to aggression—usually military action. If they speak of "Sarajevo", however, they are warning against a drift to war.[6]

From the perspective of managing China's rise, Rachman continued:

> This year's centenary of the outbreak of the First World War could do the world a great service by persuading modern politicians to spend more time thinking about Sarajevo, and less time worrying about Munich.

But this assumes the Great War was the "inadvertent" conflict of legend. In fact there was no drift, no lack of foresight, communication, realism, or calculation. There was no lack of mediation, either. In the July crisis the British government made repeated offers to mediate, but in Berlin these were seen as a complication to be neutralized. German responses were designed only to avoid public blame for the onset of war when it came.

Rather, war came in 1914, unlike in previous crises, because in that moment the great powers did not deter each other. From this perspective the reality of 1914 looks surprisingly similar to 1939, when only a credible deterrent might have stopped Hitler in his tracks. As Margaret Macmillan (2013: 503–4) has shown, the Victorians understood deterrence perfectly well. War broke out in August 1914 because, in that moment unlike all the preceding moments, the Austrian and German governments were insufficiently deterred.

The breakdown of deterrence in 1914 was fostered by rational pessimism. Pessimism changed the balance of fear: there is no option value in waiting and

[6]"Time to think more about Sarajevo, less about Munich," *Financial Times*, January 6, 2014. See also Skidelsky (2003).

no merit in avoiding risky adventures if you fear the future more than you fear your enemies. Today we should be grateful, perhaps, that China's leaders do not seem to fear the future. In today's world that fear is more likely a factor in the decisions of declining regimes in Moscow and Pyongyang.

In 1914 there was also a deeper cause of war. The war was instigated by unaccountable rulers meeting in secretive councils where the national interest was defined by military beliefs and values, and other values were excluded. In business and society, no one wanted war. The banker Max Warburg told the Kaiser: "Germany becomes stronger with every year of peace. We can only gather rewards by biding our time" (from Herwig 2011: 13). But the Kaiser did not listen.

Empirically, open political systems that aggregate widely held social preferences and exercise civilian control over military authority appear to engage in warfare more reluctantly and more selectively than their authoritarian counterparts (Levy 1988). There can be adverse side effects; for example, a democracy might be inhibited from undertaking decisive military action that would deter aggression by others. Thus a world where democracies and non-democracies coexist is not necessarily more peaceful. It seems beyond dispute, nonetheless, that if German and Austrian councils had had to listen to middle and working class opinions in 1914 there would have been no Great War.

Myth #2. Needless Slaughter

The Great War took place in an era of mass armies. This era began in the 1860s, when the railway technology first enabled the assembly and deployment of multi-million armies (Onorato et al. 2014), and ended in the 1970s when battlefield nuclear weapons and cruise missiles deprived the same mass armies of their viability.

Angell and Bloch warned that warfare in the age of mass armies would be militarily horrendous and economically and socially unbearable. In this they were partially correct: the wars were certainly horrible, but European societies proved all too capable of carrying on unprecedented war mobilizations for years at a time.

The focal point of the Great War was conflict between Germany and Russia, triggered by the gradual disintegration of the Austrian and Ottoman Empires. In fact, this was the point of the war: Germany went to war in the West, only in order to secure the conditions of victory in the East. In this sense there is an exact parallel between the two World Wars.

The two wars differ, however, in where the outcome was decided. In the Second World War, the Eastern front was decisive: this is where the main forces were concentrated and the main battles were fought. In the First World War, in contrast, the decisive engagements took place on the Western front. In fact, Germany won in the East, and was then defeated in the West.

In the West, the decisive conflict took the form of a war of attrition. The original German war plan was for an offensive across Belgium and France, ending in destruction of the French army within 6 weeks. Although the French Army suffered its worst losses of the war, it did not collapse, dug in, and was quickly buttressed by a small British expeditionary force.

Britain went to war with a strategy of attrition, which required the immediate raising of a mass army. At first there was no intention to send the new army into battle on the Western front to kill and be killed; Kitchener planned to wait, perhaps until 1917, while the French and the Russians wore the Germans down, and then intervene with "the last million men" that would decide the war (French 1982; Bourne 2005: 129).

As things turned out, however, before 1914 was over the British Army was fully engaged. The first, second, and third millions went to war long before 1917. Once the front line of trenches and dugouts settled into place, attrition became the norm. Attrition is an ugly word, designed to conceal the attempt to exchange wounds and deaths with the adversary at a favorable rate. Generals on both sides accepted casualties on a scale unthinkable by modern standards in the hope that the enemy's loss would be greater. On an average day of a war that lasted more than 4 years, more than 6000 soldiers of all nationalities died, including 2.5 thousand of the Central Powers and 3.5 thousand of the Allies (from Urlanis 1971: 209). Britain's worst day was the first on the Somme, 1 July 1916, with 20,000 killed and missing.

Attrition was a reality; was it pointless? That idea was founded on two rates of exchange: lives for lives, and lives for territory. Lives for territory: In most battles on the Western front until 1918 only a few yards changed hands for thousands or tens of thousands of casualties, and this is often counted as a measure of waste. But an emphasis on movement mistakes the purpose of combat. The purpose of combat was not to capture territory or even to achieve local breakthroughs but to destroy the enemy's fighting power. Lives for lives: When the Allied armies traded lives with the enemy, they consistently came off worse. This was a deadly problem.

For illustration, Fig. 1 reports British and German casualties in the British sector as monthly averages. It separates the 16 months from Neuve Chapelle to the days before the Somme, and the 29 months from the Somme to the Armistice. Two things are clear. First British losses increased by a multiple; at a monthly rate they were 2000 a month in the first period and 12,000 a month in the second. The main factor in this increase was the great expansion of the British forces deployed. Second, British losses consistently exceeded those of the adversary in the same sector: by nearly 2 to 1 in the first period and still by 1.5 to 1 in the second.

At that rate, the Allied policy of attrition was irrational. When Falkenhayn launched the battle of Verdun in February 1916, he expected to lose two of his own men for five French soldiers. At that rate he would drain France of blood (Herwig 1997: 182). Germany, not Britain, would have the last million.

Did the Allied generals see this? The British commanders' alleged failure to learn is embodied in Alan Clark's (1961) invented epithet: "Lions led by donkeys." Niall Ferguson (1998: 242–281) has also maintained that, given their great advantage over the Central Powers in economic capability, the Allies' failure to win the war more quickly must be explained by disorganization and incompetence: an "advantage squandered."

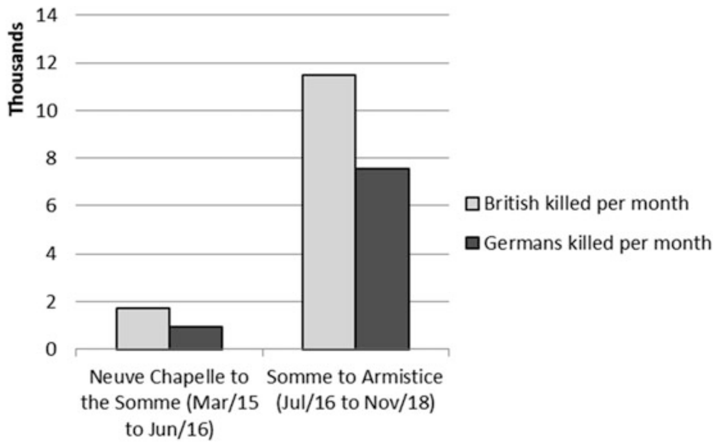


Fig. 1 British and German military deaths on the Western front, British sector, monthly average, before and after the Somme (Source: Calculated from War Office 1922: 358–362)

In fact, commanders on both sides made repeated efforts to escape the logic of attrition and avoid the slaughter. The problem was not that they did not try; the problem was that these efforts did not work.

Indirect routes to victory started with economic blockade, so Britain blockaded Germany, first at sea, and then by reducing Germany's overland trade to a minimum, while Germany blockaded Britain using submarines. But there were many countermeasures and ways around a blockade; the effects on fighting power were slow to materialize (at best) or even counterproductive and there were many countermeasures. Another indirect route seemed to lie through the Near East, where British Empire forces attacked Gallipoli, and Germany tried to ignite jihad against British colonialism; both enterprises were costly failures.

Then there was the direct route to victory, the dream of a decisive breakthrough which started always from heavier and heavier artillery bombardment of the enemy trenches, and continued (as time went by) through bombardment that became more and more accurate in timing and placement to new weapons that cascaded onto the battlefield and eventually transformed it: gas, rifle grenades and trench mortars, portable automatic weapons, and eventually tanks and airplanes. These efforts did not fail; in the end they won the war for the Allies. But they did not put an end to attrition or avoid the slaughter because the casualties only increased (Strachan 2014).

Based on manpower alone, a strategy of attrition was self-defeating: the Allies could have expected to lose the war. Kitchener's last million (men) turned out to come from America, which no one anticipated in 1914. During the war, however, the solution to attrition presented itself, and the Allies were better able to grasp this solution than the Central Powers. The stalemate on the Western front would be broken not by manpower but by firepower. Additional firepower was supplied by the Allied economies, which were much more productive than those of the Central

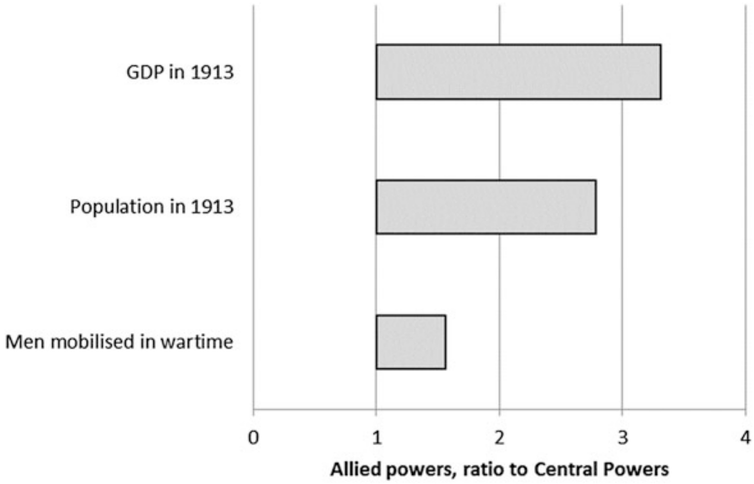


Fig. 2 The Allied material advantage: prewar GDP and population and wartime military mobilization (Sources: Prewar GDPs and populations from Broadberry and Harrison 2005a: 7–8 and 10; men mobilized from Urlanis 1971: 209)

Powers. First to grasp this was Lloyd George, who echoed Kitchener in 1915 by claiming that Britain would raise the "last million" (pounds) that would win the war (Macdonald 2006: 403).

The balance of economic advantage is easily illustrated. Figure 2 compares the Allied advantages in prewar GDP and population and in men mobilized in wartime. It is a snapshot, not a moving picture, so it rolls up Russian and American resources into a single number, ignoring the fact that Russia dropped out as America came in. Still, these were the resources available. Allied GDP was more than three times that of the Central Powers; population more than twice; men mobilized more than one and a half times. Figure 3 focuses on war production. In tanks (especially), airplanes, machine guns and rifles, the Allies comfortably out-produced the Central Powers; only in artillery did they fall short (we will see below that this margin was of doubtful benefit to the adversary).

It was in the economic dimension of attrition that the stalemate was broken, leading to Allied victory. Their economic advantage allowed the Allies to compensate for heavy casualties by superior accumulation and diversification of firepower (Prior and Wilson 1998; Strachan 2014). Allied artillery bombardments became effective against opposing trenches when intensified above a calculable threshold. New infantry weapons, combined with armor and air support, allowed the infantry to get out of the trenches and fire and move at the same time. All this was based on a vast Allied production mobilization. American reinforcement in 1917 also mattered: it added a million men (but with little experience), and some elements of firepower (but few heavy weapons) (War Office 1922: 628).

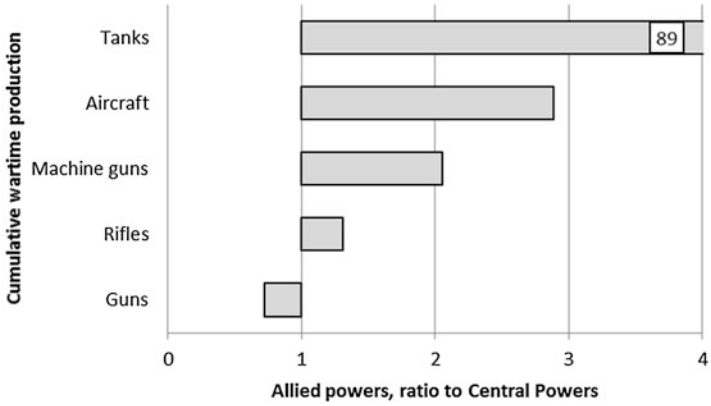


Fig. 3 The Allied material advantage: cumulative wartime production (Sources: War production from Adelman 1988: 45, except UK from Broadberry and Howlett 2005: 212, and Austria–Hungary from Schulze 2005: 88)

Mobilization of the economy for total war took time (Broadberry and Harrison 2005b). The time factor is easily illustrated. For the British the Somme is the iconic battle of the war, based on its terrible casualties and minimal gains. We will see that in July 1916, the time of this battle, the mobilization of Allied resources was still at an early stage.[7] Indeed the war after the Somme was larger in every respect. Losses were many times heavier (Fig. 1) because there were many times more soldiers who fought many times more battles (Griffith 1994: 18). The Somme ceased to be an exceptional encounter; by the end of the war, moreover, the resources being deployed in every battle exceeded those of earlier battles by large multiples. This was enabled by very large increases in combat stocks. At the time of the Somme offensive, on 9 July 1916, the British Army held more than 6.5 million shells in France and a further 1.1 million at home. By the time of the 1918 spring offensive, on 9 February, the equivalent numbers were 16.5 million and 11.3 million (War Office 1922: 481).

Behind this lay the expansion of production, illustrated in Fig. 4. In the 30 months from the Somme to the end of 1918, British industry produced rifles at twice the rate, shells at 5 times, machine guns and aircraft at 9 times, and tanks at 33 times the rate of the similar period up to the Somme.

Not surprisingly, there was an answering mobilization on the German side, most notably the "Hindenburg plan" of war-industry construction adopted in August

[7]Griffith (1994) has called the period after the Somme the "larger second half of the war." He criticizes much of the British history and literature of the war as suffering from an "early war bias": it has typified the conduct of the war, based on the experience of the early months of "amateurism, blundering, and fumbling" when most of the serious fighting was done by the French. At this time most of the English poetry was written (and perhaps the poets were able to write it because the British sector was quiet most of the time).

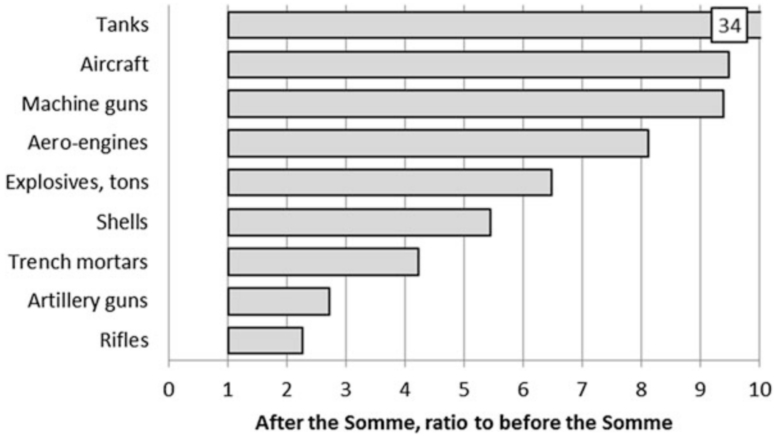


Fig. 4 British war production after the Somme, monthly average, ratio to before the Somme (Source: Annual data from Broadberry and Howlett 2005: 212. Figures for 1916 are distributed equally between the first and second halves of the year)

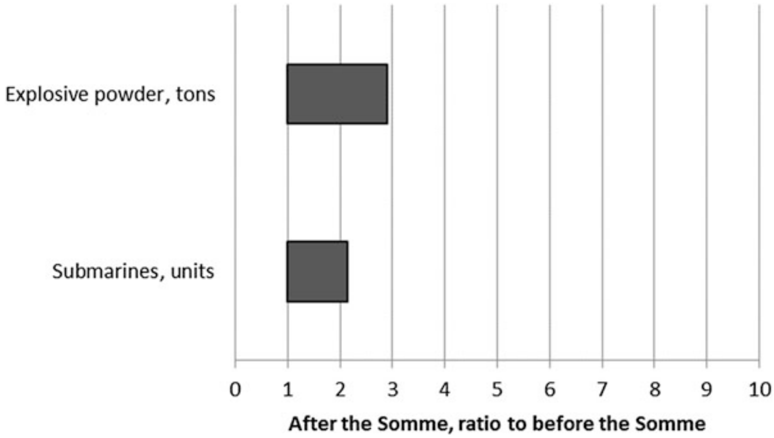


Fig. 5 German war production after the Somme, monthly average, ratio to before the Somme (Sources: Explosive powder: Feldman 1966: 152–3, 268, 272, 494 gives figures for December 1915 (4000 tons), July 1916 (6000 tons), February 1917 (6400 tons), April 1917 (8000 tons), May 1917 (9200 tons), and April 1918 (12,000 tons). I assume that these figures were selected because they are salient, and I interpolate linearly between them. I take 1000 tons as the monthly figure for August 1914, and I assume that a monthly output of 12,000 tons (the target of the Hindenburg plan) was maintained through November 1918. Submarines: Davis and Engerman 2006: 232–233)

1916. The best account, almost half a century old, and with many gaps, is still that of Feldman (1966). The Hindenburg plan was focused on guns and explosive powder. Figure 5 infers that the monthly rate of powder production in Germany from the Somme to the Armistice was three times that of the first half of the war (the horizontal axis of this figure is scaled to match that of Fig. 4). As for guns, it is known (Fig. 3) that over the war period as a whole Germany alone outproduced the Allies (including both Russia and America), but we do not know the timing of this accomplishment. Figure 5 also reports Germany's submarine construction after the Somme at around twice the rate beforehand. These numbers, while hardly spectacular, are evidence of German production success to compare with that of the Allies.

Two things undermine the German economic record, however. First, the German production program was insufficiently diversified: it neglected the airplanes and tanks that could give supporting cover to the attacking infantry and would prove decisive in 1918.

Second, while the Hindenburg plan was too narrow, the huge efforts that it forced were also too much for the German economy. A textbook description of unsustainable mobilization is provided by the memoir of German Interior Minister Karl Helfferich (cited by Feldman 1966: 273), according to whom, by mid-1917, there were:

> Everywhere half-finished and finished factories that cannot produce because there is no coal and there are no workers available. Coal and iron were expended for these constructions, and the result is that munitions production would be greater today if no monster programme had been set up but rather production had been demanded according to the capacities of those factories already existing.

The excessive mobilization precipitated the disintegration of the German economy, the collapse of living standards, and an urban famine. As a result Germany could not reap the dividend from its victory on the Eastern front. By 1918 more than two million men available for military service were being held back from the front to work in Germany, alongside thousands of soldiers returned from the front (Bessel 1988: 24–25). By now the Germany army had too many guns for the men available to fight, and still not enough food (Herwig 1997: 264, 410). On one hand "the new artillery, trench mortars, and machine guns rusted on loading docks." On the other hand, the troops were afflicted by "hunger and thirst," which drove them through the year's spring offensive: as they advanced they quickly exhausted their own supplies, which they made up by looting the abundance that they found in Allied stockpiles, but the search for food and wine slowed the advance.

In short, attrition worked, not in the generals' understanding of the term which was indeed narrow-minded and pre-modern (Offer 1989: 352), but in Lloyd George's modernized interpretation: When the financial and industrial strength of the central powers was finally exhausted, the Allies still had the last million.

Should the Allied victory have come sooner? Ferguson (1998) argued that, given their material advantage, the Allies should have won in 2 years. Evidently "the Germans were significantly better at mobilizing their economy for war than the Western powers." The Allied advantage, he concluded, was squandered. But this is not a reasonable conclusion. Mistakes were made on both sides. The British failed to learn in time from previous industrialized wars (Trebilcock 1975) and from early battles on the Western front (Prior and Wilson 1998; see also Griffith 1994: 192–200). But the German generals that ran the domestic war effort from 1916 also made terrible mistakes. They started from a position of material inferiority. They ended up with war factories they could not supply or operate and soldiers that they could not feed. These were mistakes they could not afford.

To conclude, in the Great War, despite material advantage, the Allies could not escape the war of attrition. Attrition began with mass killing. The scale of mass killing cannot be described as efficient; inhumanity was mixed with disastrous miscalculations and failures to learn. But this was war, and mass killing was the result of warfare in the era of mass armies, not of the particular form warfare took at this time. At the same time attrition was not only mass killing. There was also economic attrition. It was the combined attrition in both economic and military dimensions that defeated the Central Powers.

Myth #3. The Food Weapon

Food was an essential element of two world wars (Collingham 2010). Moreover, food security was a core element of German war preparations (Lee 1975). Despite such preparations, many believed, Germany was strangled by the British (later Allied) blockade. The food weapon appeared to have been decisive: Germany was starved into submission.

This belief has historic significance. After the war it helped to sustain the notion (attributed to Germany's wartime leaders Hindenburg and Ludendorff) that Germany remained unbeaten militarily; the army was betrayed by the surrender of the home front. The memory of the blockade also ran deep in the National Socialists' project to restructure Europe in Germany's interest by force, as when Hitler (cited by Collingham 2010: 37) remarked in 1939: "I need the Ukraine, so that no one is able to starve us again, like in the last war."

The idea that Germany was starved into defeat would have astonished prewar observers. The British and German prewar diets were quite comparable. At the outbreak of war Germany imported only 20–25 % of calories for human consumption; for Britain the equivalent number was 60 %. It was natural for Angell and Bloch to suppose that in wartime, British consumers would starve first. Yet all measures of wartime trends show a contrast that was unfavorable to the German consumer. During the war British food supplies were somewhat constrained and their average composition deteriorated; in 1918, the average household consumed more bread, less fat, and substantially less meat than in 1913. In Germany, in contrast, in 1918 the average household ate less of everything, and supplies of meat and fats had collapsed.

Germany also compares unfavorably in food distribution. There, it was families on lower incomes that were less protected from average trends. In Britain the access of poorer families to food improved relative to the average (Gazeley and Newall 2013); this is more likely attributable to the high demand for all kinds of labor than to rationing, which was introduced only at the end of 1917 (for sugar) or during 1918 (for some meats and fats). In Germany price ceilings and rationing came in 1916 and covered bread and flour, meat, fats, and oil. But rations supplied little more than half of required calories, so everyone had to find unofficial sources to survive. In this setting the wealthy had the advantage. Nutritional deprivation has been observed in the heights of soldiers and in the heights and weights of schoolchildren born before and during World War I (Blum 2013; Cox 2014). Both show average declines and increases in inequality.

Finally, excess mortality among German civilians wartime is put at around 750,000, most likely because of hunger and hunger-related disease (Davis and Engerman 2006: 204).

The blockade was the adversary's salient intervention in Germany's food supplies, and it is easy to leap to the conclusion that the blockade was therefore the cause of German hunger.[8] But this story is confounded by two factors. One is straightforward: Germany chose to go to war with its principal trading partners. Angell and Bloch had argued forcefully that great powers heavily dependent on trade should not attack the sources of their own prosperity. But this is exactly what Germany did (and Shinzo Abe was right to note the fact). The German economy was much more interlinked with its future adversaries than its future allies. In 1913, Britain, France, Italy, and Russia accounted for 36 % of prewar German trade (Gartzke and Lupu 2012: 131).[9] The same figure for Austria–Hungary, Bulgaria, and the Ottoman Empire was only 12 %. From this perspective, Alan Kramer (2013) has pointed out, much of the "blockade" was no more than an Allied decision not to supply the enemy across the front line.

Another confounding factor is suggested by the fact that the loss of trade was not the only supply shock disturbing the wartime food market. Prewar plans for wartime autarky assumed that German farmers would farm more intensively to feed the nation (Lee 1975). But the opposite came about, because war mobilization stripped German farms of young men, horses, and nitrates. War mobilization also diverted domestic industries from producing the manufactured goods that farmers needed to supplying weapons for the front line. As food prices soared, farmers retreated into self-sufficiency. When civilian officials stepped in to control prices, the farmers' aversion to trade only increased.

In his economic history of the war Gerd Hardach (1987: 34) asked how the blockade interacted with Germany's economic mobilization. He conjectured:

> The tremendous economic decline of the Central Powers between 1914 and 1918 was caused less by the blockade than by the excessive demands made on their economies by the war.

[8]This leap is illustrated by Mary Cox's (2014) title: "Hunger games: or how the Allied blockade in the First World War deprived German children of nutrition, and Allied food aid subsequently saved them." Nothing in her otherwise insightful and important article proves a direct channel of causation from blockade to hunger.

[9]Ferguson (1998: 253) gives a higher figure for the share of Germany's prewar imports from wartime adversaries: 48 %.

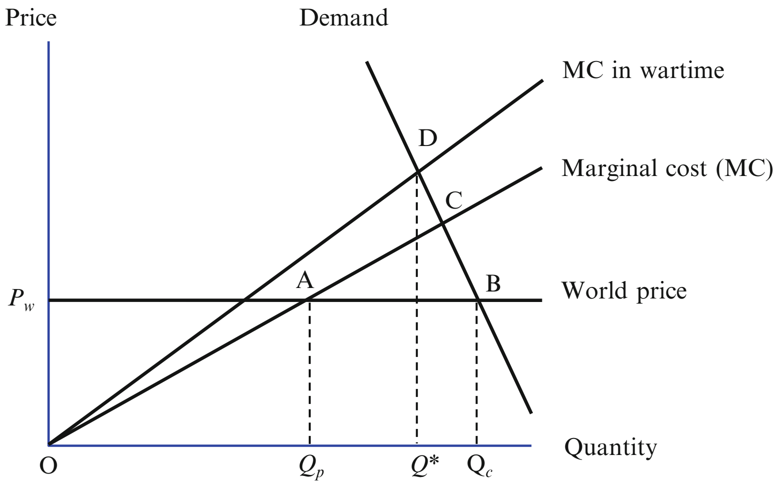


Fig. 6 The Hardach conjecture: the German market for calories. Note: The "blockade effect" is shown by the *triangle ABC*, and the "mobilization effect" by the *triangle OCD*

Hardach did not suggest how to implement this comparison in the German market; here is a simple way to think about it. Start from the fact that before the war Germany imported at most one quarter of calories for human consumption, producing the other three quarters on its own territory. In that context the war induced two welfare losses. One, arising from wartime obstacles to external trade, raised the costs of the one quarter of calories that was previously imported. The other, arising from wartime mobilization, raised the costs of the three quarters produced at home. Is it reasonable to suppose that the loss associated with the one quarter was larger than the loss associated with the three quarters?

Figure 6 illustrates the point. P_w is the peacetime world price of calories; Q_p is peacetime calories produced and Q_c is calories consumed, the gap being filled by imports at the world price. Suppose that war cuts off all trade—an overstatement of the case. The welfare loss from the blockade is the triangle ABC. Suppose that at the same time war mobilization raises the costs of domestic production. Then the welfare loss from mobilization is the triangle OCD. While the height of each triangle cannot be ascertained, its base is known. The trade loss is proportional to prewar trade, whereas the mobilization loss is proportional to prewar consumption. Since the share of prewar trade in consumption was at most one quarter (and not all trade was cut off), we can reasonably presume that the welfare loss from mobilization was greater than the loss from the blockade.[10] Further welfare losses could have arisen from price ceilings and rationing. On the evidence already cited they redistributed welfare adversely but are ignored in the figure.

[10]This line of argument has a precedent in Williamson (1968: 21–23).

Wartime mobilization ended well before the lifting of the blockade, which was maintained after the Armistice and, extended to the Baltic, became even tighter. Until Germany's acceptance of the Treaty of Versailles in June 1919, with the fighting over and German soil under German control, trade sanctions were the Allies' only coercive lever to ensure that Germany came to terms. Among Germans the continuation of the blockade was bitterly unpopular and became a source of lasting resentment. Yet, as Offer (1989: 388–391) reports, prices did not rise and rations did not fall. One explanation is that the end of wartime mobilisation compensated for the intensified blockade.

It was both plausible and convenient for politicians of the war period and later to blame German's wartime economic difficulties on the Allied blockade. This must be largely a myth. The blockade was not the only factor in the disruption of German trade. The disruption of trade was not the only factor that disrupted the German internal market for food. Arguably, the military mobilization of agricultural resources into war, and the economic mobilization of industry, had a larger disruptive effect than the shock from foreign trade.

Myth #4. Folly at Versailles

The Treaty of Versailles of 1919 and the Reparations Commission that it established imposed heavy burdens on Germany. Having witnessed the negotiations Keynes (1920) condemned the outcomes on two grounds: they violated the terms of the Armistice (which limited German reparations to making good civilian damages arising from the war) and the resulting burden on the German economy was intolerable and would be counterproductive. The alleged repercussions could not have been more serious. According to the financier and philanthropist George Soros (2014), for example, the French "insistence on reparations led to the rise of Hitler." There are present-day implications for, Soros continues, "Angela Merkel's [similar] policies are giving rise to extremist movements in the rest of Europe."

Another supposed channel of causation is the German hyperinflation of 1923. According to Brookings Vice-President (formerly a Turkish government minister and UN administrator) Kemal Derviş (2014), "had Germany's hyperinflation of the 1920s—a direct result of the war—been avoided, Hitler [might] well never have risen to power."

On a sober evaluation the burden of German reparations determined in 1921 was certainly heavy and probably unwisely so. The evidence is plain to see in the better outcome of 1945, when the victors based retribution more on personal culpability than collective responsibility. Still, the mistakes of 1919–1921 should be seen in broader perspective.

In 1921 the Reparations Commission issued A, B, and C bonds. The C bonds were issued as a symbol to appease various Allied constituencies, and their redemption was deferred indefinitely; it was openly acknowledged that only the A and B bonds were within Germany's ability to pay. The A and B bonds together have been valued at 1.25 years of German GDP in 1921 (Eichengreen 1995: 129). If we add Germany's ordinary public debt, which Ritschl (2005: 69) puts at half of GDP in 1920, then we arrive at around 1.75 years of German GDP. This looks heavy but not overwhelming when compared with the sovereign debt liabilities of France and the United Kingdom in 1921 (2.6 and 1.5 times GDP respectively).[11]

Much has been made of the so-called transfer problem arising from the burden of current repayments. Taking into account Germany's Treaty losses of territory and capacity, and adding non-reparations obligations to the Allies, Webb (1986) estimated the current burden at around 10 % of Germany's postwar GDP. The implied strain on the balance of payments would have been severe. It was expected, however, that Germany would smooth out repayments by commercial borrowing.

Germany's center-left government did not want to smooth out reparations, however; it did not want to cover them at all. At this point the problem ceased to be the burden of attempts to comply, for Germany made few attempts (Marks 1978). The clash with the Allies led, in early 1923, to the occupation of the Ruhr as the French tried to extract reparations by compulsion. Now the German government wrote another blank check, this time to fund the efforts of the local population to frustrate collection. As Webb (1986) has described, the anticipation of unbounded future fiscal deficits triggered a disastrous hyperinflation.

In Germany, therefore, reparations led to hyperinflation, but through a channel that was political as much as economic. Moreover, while reparations contributed to hyperinflation in the German case, they were not a necessary condition elsewhere. For there were simultaneous hyperinflations across the region from Austria and Hungary to Poland, the Baltic, and Russia (Bresciani–Turroni 1937; Dornbusch 1991). As Keynes (1920: 223) presciently remarked, "The inflationism of the currency systems of Europe has proceeded to extraordinary lengths. The various belligerent Governments, unable, or too timid or too short-sighted to secure from loans or taxes the resources they required, have printed notes for the balance." Regardless of their new and fragile constitutions, governments across central and eastern Europe spent beyond their means, and compliant central bankers monetized the resulting debts.

To generalize, every hyperinflation of the period began from a civil war of attrition (in the political-economy sense of Alesina and Drazen 1991) among the various classes of bondholder, debtor, and taxpayer, each of which preferred to postpone stabilization in the hope of shifting the burdens of stabilization onto others. The German meltdown was unique only in the role of the Allies as external bondholders.

After stabilization came the first of many reschedulings (Schuker 1988). There was also a peace dividend. Hantke and Spoerer (2010) note something that Keynes and others entirely neglected: the Treaty provisions that limited German interwar rearmament gave Germany fiscal breathing space. Restrictions on the size and

[11]Data from http://www.reinhartandrogoff.com/data/ (accessed 23 February 2014), described by Reinhart and Rogoff (2011).

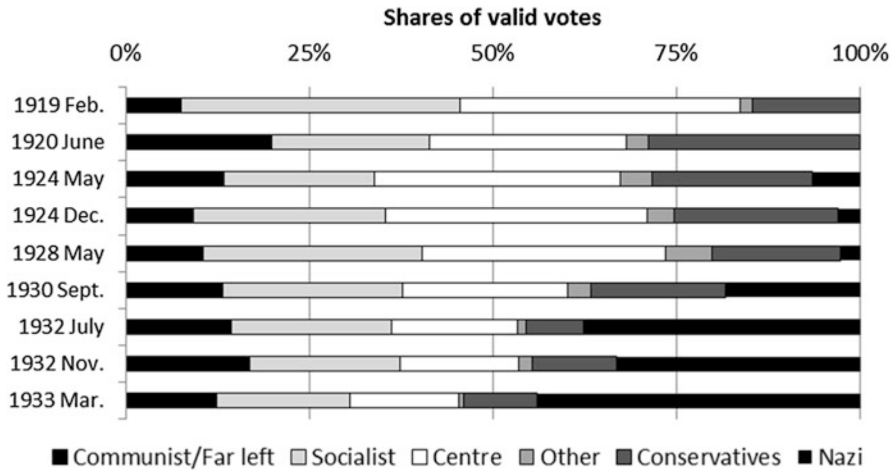


Fig. 7 Elections to the Reichstag, February 1919 to March 1933 (Sources: Figures for 1920–1933 from Berghahn 1982: 284–285, supplemented by February 1919 from "German Federal Election, 1919" at http://en.wikipedia.org/wiki/German_federal_election,_1919 (accessed 23 February 2014). Votes for the German People's Party (DVP) under Stresemann are counted with the center from 1923 to 1929, and with the conservative parties otherwise)

equipment of Germany's armed forces reduced the burden of military spending. In equilibrium they probably also reduced military spending across Europe. To estimate the peace dividend requires a counterfactual hypothesis which is not straightforward, but in Hantke and Spoerer's most conservative scenario it was large enough to cover at least 90 % of the reparations actually paid in the years from 1924 to 1929.

If the economic implications of the Treaty have been oversold, the same is true of its political consequences. The electoral history of the Weimar Republic may conveniently be broken into three phases: 1919 to May 1924, with three parliamentary elections), a period that included the imposition of the Treaty of Versailles, the announcement of reparations, and the hyperinflation; December 1924–1928, marked by two elections, the first preceded by announcement of the Dawes Plan to reschedule Germany's debts and the second being the last before the Great Depression; and 1930–1933 (four elections, the first coming just after the second rescheduling of the Young Plan, and the last ending in the Hitler dictatorship.

Figure 7 shows the evolution of votes for the German parliament through the life of the Weimar republic, from the first election of February 1919 to the last of March 1933 that put Hitler into power. The figure arranges Germany's parties from far left to far right. The parties of the far left aimed to overthrow the new parliamentary democracy in favor of a Soviet republic. The conservative right pressed for a return to monarchy. A key fraction therefore was the share of support for the democratic parties of the center and moderate left.

In the election of February 1919 the democratic parties swept the board. The imposition of the Versailles Treaty in the summer of that year hit them hard—although one must also take into account that some degree of disillusionment with the peace was inevitable. With the scale of reparations not yet fixed, in June 1920 the democratic parties lost 30 % of their former support. The lost votes went to the extremes of the far left and right. This was a dangerous moment, because the supporters of the constitution were briefly in a minority.

The June 1920 election was followed by the announcement of reparations (in 1921) and the subsequent turmoil that ended in the 1923 hyperinflation. In the May 1924 elections, the national socialists put in a first showing, winning nearly 7 % of seats. Yet their gain was entirely at the expense of the conservatives, who were against democracy anyway. At the same time the communists lost ground. The democratic parties recovered their majority. At a second election held in December 1924 the extremists' vote collapsed, the national socialists almost disappeared, and the democratic parties gained more ground.

In May 1928, three and a half years later, despite continued agitation on the reparations issue, the trend towards democratic consolidation continued. The communists gained little ground and support for the national socialists remained below 3 %.

The economic hammer blow of the Great Depression fell in the summer of 1929. Unemployment among German workers rose year by year to a peak of 30 % in the summer of 1932. German farmers' prices and incomes also collapsed. It was only now that the far right broke through to national significance and then to power (Van Riel and Schram 1993; King et al. 2008).

In other words, from the Dawes Plan to the Great Depression the German electorate showed a substantial and growing majority for constitutional rule by democratic parties. Neither the Treaty of Versailles, nor the announcement of reparations, nor the initial conflict over payments, nor the hyperinflation that followed, show any persistent effects on German politics. Were it not for the Great Depression, Hitler and his infamous co-conspirators would have lived to the 1960s and died in obscurity in their beds.

In setting out to punish Germany for the war the authors of the Treaty of Versailles lacked wisdom and enlightened self-interest. Contemporary critics of the Treaty in Germany and abroad made the most of this. But it is wrong to look here for the causes of the Hitler dictatorship and World War II. There were plenty of precursors of a future conflict before the Treaty was ever signed. Germany was still fighting in September 1918 when the high command launched its Fatherland Party demanding peace with annexations East and West (Howard 2002: 98). Germany was not only fighting but winning in September 1914 when Bethmann Holweg advanced his program to restructure Europe as a German colonial empire (Hastings 2013: 100). It was the outcome of the war, not the terms of the peace, that Germany's far right would not accept. What is striking is how muffled and insignificant their voice became during the 1920s while most other Germans worked at becoming a normal country.

The most that can be held against the Treaty of Versailles is that it did not help. It was bad diplomacy and, given what diplomacy is supposed to achieve, that is bad enough.

Concluding Remarks

We have reviewed four widespread narratives of the Great War. Each is at the crossroads of economics, politics, and strategy. Myths are not necessarily baseless, and we have tried to distinguish truthful elements, but their part in the overall story was often relatively small.

A myth of the war's origin is that it was an inadvertent conflict that transpired without intention or calculation. In fact, the decisions that led to the Great War were calculated with considerable foresight of the wider costs and consequences. The spirit of those that started the war is usefully defined as "rational pessimism."

A myth surrounding the waging of the war sees the fighting on the Western front as needless attrition. There was attrition, and attrition was pursued deliberately on both sides. From the Allied standpoint this looks scarcely rational because the rate of exchange of casualties was always adverse. The missing dimension was economic: the "last million" was measured in productive resources, not only manpower. The Allies outproduced the Central Powers in firepower (and everything else) and this was the basis of victory. There was no escape from attrition, and no other way of winning the war.

A myth surrounds the end of the war: that Germany was starved out by the food weapon. It is true that German civilians suffered greatly, and that the Allied blockade contributed, but it seems likely that German actions contributed more. These included the decision to attack Germany's main trading partners and the impact of Germany's economic mobilization on the internal food market.

A final myth surrounds the peace: that the Treaty of Versailles, which concluded this war, laid the foundations for the next one. Many aspects of the Treaty would seem to fail the test of enlightened self-interest on the Allied side. Despite this, the electoral impact of the Treaty, the reparations issue, and the hyperinflation that followed was short-lived. For most of the 1920s, German society was set on a course of democratic consolidation.

The Great Depression, which struck Germany in 1929, brought back to life dark forces of violent nationalism. These forces were engendered far back in German history. The war let them loose; Germany's defeat put them back in a cage. Weimar democracy put them into a coma. It was the Great Depression that revived them, so that they sprang back into life.

References

Adelman, J.R. 1988. *Prelude to the cold war: The Tsarist, Soviet, and U.S. armies in two World Wars*. Boulder: Lynne Riener.

Alesina, A., and A. Drazen. 1991. Why are stabilizations delayed? *American Economic Review* 81(5): 1170–1188.

Angell, N. 1910. *The great illusion: A study of the relation of military power to national advantage*. London: Heinemann.

Berghahn, V.R. 1973. *Germany and the approach of war in 1914*. London: Macmillan.

Berghahn, V.R. 1982. *Modern Germany: Society, economy and politics in the twentieth century*. Cambridge: Cambridge University Press.

Berghahn, V.R. 2013. Origins. In *The Cambridge history of the First World War*, vol. 1, ed. J. Winter, 16–38. Cambridge: Cambridge University Press.

Bessel, R. 1988. The Great War in German memory: The soldiers of the First World War, demobilization, and Weimar political culture. *German History* 6(1): 20–24.

Bloch, I.S. 1899. *Is war now impossible?* London: Grant Richards.

Blum, M. 2013. War, food rationing, and socioeconomic inequality in Germany during the First World War. *Economic History Review* 66(4): 1063–1083.

Bourne, J. 2005. Total war I: The Great War. In *The Oxford history of modern war*, 2nd ed, ed. C. Townshend, 117–137. Oxford: Oxford University Press.

Bresciani–Turroni, C. 1937. *The economics of inflation*. London: Routledge.

Broadberry, S., and M. Harrison. 2005a. The economics of World War I: An overview. In *The economics of World War I*, ed. S. Broadberry and M. Harrison, 3–40. Cambridge: Cambridge University Press.

Broadberry, S., and M. Harrison. 2005b. *The economics of World War I: A comparative quantitative analysis*. Paper presented at the Economic History Association, Toronto, 16–18 September 2005.

Broadberry, S., and P. Howlett. 2005. The United Kingdom during World War I: Business as usual? In *The economics of World War I*, ed. S. Broadberry and M. Harrison, 206–234. Cambridge: Cambridge University Press.

Clark, A. 1961. *The donkeys*. London: Hutchinson.

Clark, C. 2013. *The sleepwalkers: How Europe went to war in 1914*. London: Penguin.

Collingham, L. 2010. *The taste of War: World War Two and the battle for food*. London: Allen Lane.

Cox, M.E. 2014. Hunger games: Or how the allied blockade in the First World War deprived German children of nutrition and allied food aid subsequently saved them. Forthcoming in The Economic History Review.

Davis, L.E., and S.L. Engerman. 2006. *Naval blockades in peace and war: An economic history since 1750*. Cambridge: Cambridge University Press.

Derviş, K. 2014. The Great War and global governance. Project Syndicate, 8 January. Available via http://www.project-syndicate.org/commentary/kemal-dervi--hopes-that-this-year-s-centennial-of-the-start-of-world-war-i-will-lead-to-greater-reflection-on-ways-to-minimize-risk-at-an-acceptable-cost. Accessed 8 Jan 2014.

Dornbusch, R. 1991. Experiences with extreme monetary instability. In *Managing inflation in socialist economies in transition*, ed. S. Commander, 179–180. Washington, DC: The World Bank.

Eichengreen, B. 1995. *Golden fetters: The gold standard and the great depression, 1919–1939*. Oxford: Oxford University Press.

Erdmann, K.D. (ed.) 1972. Kurt Riezler. Tagebücher, Aufsätze, Dokumente. Vandenhoeck and Ruprecht, Göttingen, 181–193 abridged and translated by Patch W L at http://home.wlu.edu/~patchw/His_224/1914.htm. Accessed 10 Mar 2013.

Feldman, G.D. 1966. *Army, industry, and labor in Germany, 1914–1918*. Princeton: Princeton University Press.

Ferguson, N. 1998. *The pity of war*. London: Allen Lane.
Ferguson, N. 2005. *1914: Why the world went to war*. London: Penguin.
French, D. 1988. The meaning of attrition, 1914–1916. *English Historical Review* 103(407): 385–405.
Fromkin, D. 2007. *Europe's last summer: Who started the Great War in 1914?* New York: Doubleday.
Gartzke, E., and Y. Lupu. 2012. Trading on preconceptions: Why World War I was not a failure of economic interdependence. *International Security* 36(4): 115–150.
Gazeley, I., and A. Newall. 2013. The First World War and working-class food consumption in Britain. *European Review of Economic History* 17(1): 71–94.
Griffith, P. 1994. *Battle tactics of the western front: The British Army's art of attack, 1916–1918*. New Haven: Yale University Press.
Hamilton, R.F., and H.H. Herwig. 2004. On the origins of the catastrophe. In *Decisions for war, 1914–1917*, ed. R.F. Hamilton and H.H. Herwig, 225–252. Cambridge: Cambridge University Press.
Hantke, M., and M. Spoerer. 2010. The imposed gift of Versailles: The fiscal effects of restricting the size of Germany's armed forces, 1924–9. *Economic History Review* 63(4): 849–864.
Hardach, G. 1981. *The First World War, 1914–1918*. Berkeley: University of California Press.
Hardach, G. 1987. *The First World War, 1914–1918*. Harmondsworth: Penguin.
Hastings, M. 2013. *Catastrophe: Europe goes to war, 1914*. London: William Collins.
Herwig, H.H. 1987. Clio deceived: Patriotic self-censorship in Germany after the Great War. *International Security* 12(2): 5–44.
Herwig, H.H. 1997. *The First World War: Germany and Austria–Hungary, 1914–1918*. London: Arnold.
Herwig, H.H. 2002. Germany and the "short war" illusion: Towards a new interpretation? *Journal of Military History* 66(3): 681–693.
Herwig, H.H. 2011. "Military doomsday machine"? The decisions for war 1914. *Journal of Military and Strategic Studies* 13(4): 1–15.
Howard, M. 2002. *The First World War: A very short introduction*. Oxford: Oxford University Press.
Keynes, J.M. 1920. *The economic consequences of the peace*. London: Macmillan.
King, G., O. Rosen, M. Tanner, and A.F. Wagner. 2008. Ordinary economic voting behavior in the extraordinary election of Adolf Hitler. *Journal of Economic History* 68(4): 951–994.
Kramer, A. 2013. Blockade and economic warfare. In *The Cambridge history of the First World War*, vol. 2, ed. J. Winter, 460–490. Cambridge: Cambridge University Press.
Lee, J. 1975. Administrators and agriculture: Aspects of German agricultural policy in the First World War. In *War and economic development: Essays in memory of David Joslin*, ed. J.M. Winter, 229–238. Cambridge: Cambridge University Press.
Levy, J.S. 1988. Domestic politics and war. *Journal of Interdisciplinary History* 18(4): 653–673.
Lloyd George, D. 1938. *War memoirs*, vol. 1. London: Odhams.
Macdonald, J. 2006. *A free nation deep in debt: The financial roots of democracy*. Princeton: Princeton University Press.
MacMillan, M. 2013. *The war that ended peace: How Europe abandoned peace for the First World War*. London: Profile.
Marks, S. 1978. The myths of reparations. *Central European History* 11(3): 231–255.
McMeekin, S. 2011. *The Russian origins of the First World War*. Cambridge, MA: Belknap.
Offer, A. 1989. *The First World War: An agrarian interpretation*. Oxford: Clarendon.
Onorato, M.G., K. Scheve, and D. Stasavage. 2014. Technology and the era of the mass army. *Journal of Economic History* 74(2): 449–481.
Prior, R., and T. Wilson. 1998. *Command on the western front. The military career of Sir Henry Rawlinson, 1914–1918*. Barnsley: Pen & Sword.
Purcell, H. 2006. *Lloyd George*. London: Haus.
Reinhart, C.M., and K.S. Rogoff. 2011. From financial crash to debt crisis. *American Economic Review* 101(5): 1676–1706.

Ritschl, A. 2005. The pity of peace: Germany's economy at war, 1914–1918 and beyond. In *The economics of World War I*, ed. S. Broadberry and M. Harrison, 41–76. Cambridge: Cambridge University Press.

Schelling, T.C. 1966. *Arms and influence*. New Haven: Yale University Press.

Schuker, S.A. 1988. *American "reparations" to Germany, 1919–33: Implications for the Third-World debt crisis*. Princeton: Princeton Studies in International Finance.

Schulze, M.S. 2005. Austria–Hungary's economy in World War I. In *The economics of World War I*, ed. S. Broadberry and M. Harrison, 77–111. Cambridge: Cambridge University Press.

Skidelsky, R. 2003. *The origins and consequences of the First World War*. Brighton College Lecture, Tuesday, November 25.

Soros, G. 2014. The world economy's shifting challenges. Project Syndicate, 2 January. Available via http://www.project-syndicate.org/commentary/george-soros-maps-the-terrain-of-a-global-economy-that-is-increasingly-shaped-by-china. Accessed 2 Jan 2014.

Strachan, H. 2014. *The First World War*, Revised ed. London: Simon & Schuster.

Trebilcock, C. 1975. War and the failure of industrial mobilization: 1899 and 1914. In *War and economic development: Essays in memory of David Joslin*, ed. J.M. Winter, 139–164. Cambridge: Cambridge University Press.

Urlanis, B. 1971. *Wars and population*. Moscow: Progress.

Van Riel, A., and A. Schram. 1993. Weimar economic decline, Nazi economic recovery, and the stabilization of political dictatorship. *Journal of Economic History* 53(1): 71–105.

War Office. 1922. *Statistics of the military effort of the British Empire during the Great War, 1914–1920*. London: HMSO.

Webb, S.B. 1986. Fiscal news and inflationary expectations in Germany after World War I. *Journal of Economic History* 46(3): 769–794.

Williamson, O.E. 1968. Economies as an antitrust defense: The welfare tradeoffs. *American Economic Review* 58(1): 18–36.

War and Inflation in the United States from the Revolution to the Persian Gulf War

Hugh Rockoff

Abstract The institutional arrangements governing the creation of new money in the United States have changed dramatically since the Revolution. Yet beneath the surface the story of wartime money creation has remained much the same. During wars against minor powers, the government was able to fund the war by borrowing and levying taxes. During wars against major powers, however, the story was very different. In major wars there came a point when further increases in taxes could not be undertaken for administrative or political reasons, and further increases in borrowing could not be undertaken except at higher interest rates, rates that exceeded what was considered fair based on prewar norms. At those moments governments turned to the printing press. The result was substantial inflation.

Keywords Interest • Inflation • Borrowing • Adam Smith • John Stuart Mill • Robert Barro • John Maynard Keynes • [American] Revolution • War of 1812 • (American) Civil War • National Banking Act • U.S. Treasury • Federal Reserve • Benjamin Franklin • First Gulf War • Farley Grubb • Continental dollar • Virginia dollar • Continental Congress • Nathaniel Greene • Charles Lee • Daniel Morgan • French and Indian War • Congress • Debt • Milton Friedman • Anna Schwartz • Mexican War • Greenback • Supreme Court • John Sherman • Thaddeus Stevens • Jay Cooke • Homestead Act • Spanish-American War • Philippine-American War • World War I • World War II • Bonds • Korean War • Korean War Accord • Vietnam War • William McChesney Martin • Lyndon B Johson

Money Creation as a Means of War Finance

When the United States went to war against major powers it resorted to the printing press to help finance the war. In every case the result was a substantial inflation. In wars against second-tier powers, however, the United States, for the most part,

H. Rockoff (✉)
Department of Economics, Rutgers University, 75 Hamilton Street,
New Brunswick, NJ 08901, USA
e-mail: Rockoff@econ.rutgers.edu

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*,
Studies in Economic History, DOI 10.1007/978-981-10-1605-9_7

avoided the printing press and inflation.[1] At one level, the explanation is simple. A war against a major power requires more resources than a war against a second-tier power, and forces the government to dig deeper into its bag of tricks to finance the war. There is, however, a bit more to say about how, why, and with what consequences governments have turned to the printing press.

A major war means that there will be a large, although temporary, need for additional revenues. Borrowing is the obvious first choice. But if borrowing proceeds too far, at some point nominal interest rates will rise. For several reasons to be discussed below, war governments were loath to allow nominal interest rates to rise above prewar norms. Tax increases that were sufficient to finance future interest and principal payments helped to make borrowing credible and help maintain interest rates. But taxes were difficult to increase quickly, and if raised too far too fast had the potential to arouse opposition and undermine support for the war. Faced with the prospect of breaching prewar interest rate norms, and unable or unwilling to raise taxes sufficiently to finance the war, governments turned to the printing press. Although a simple story, it does, as I will show, fit the facts.

There were several reasons why war governments were unwilling to see prewar interest rate norms breached. (1) The analogy with personal or business finance undoubtedly was influential. If I can afford a mortgage at the same rate as other homebuyers, I am making a sound investment; if I can only get a mortgage at a subprime rate I have a clear sign from the market that I am overextended. Surely, the argument goes, what is true for an individual, or a family, or a business must be true for the government. If the interest rate my government must pay rises it must be overextended; it must have done something wrong. (2) Although many citizens may be unaware of the amounts being borrowed, some will be disturbed by rising indebtedness and the prospect of higher taxes in the future, a fear likely to be amplified by rising interest rates. (3) Rising nominal rates may be interpreted as a sign of economic weakness, undermining public confidence in the war effort, discouraging allies, and encouraging enemies. (4) The most important reason, however, as will be clear as we exam the statements of government officials below, was simply the ancient prejudice against usury.

Typically, the common man or woman, and the politicians who represent them, favored the borrower over the lender; the victim over the shylock. Higher interest rates on government debt meant that the taxpayer was being victimized by the shylocks. If borrowing by the government raised rates in general, then even people borrowing for private purposes were victimized. Printing money and using it to support the price of government bonds appeared to be only fair. True, printing money might cause inflation. But inflation hurt lenders and helped borrowers; hurt the shylocks and helped their victims. Higher rates in wartime were considered especially loathsome. War was a time when the nation was pulling together for

[1] The one partial exception is the Vietnam War. The war, as discussed below, did contribute to the excessive monetary growth that characterized the late 1960s and early 1970s. But there were other factors besides the war.

a common purpose. Lenders should not be allowed to profit from the special circumstances in wartime. In this respect the decision to suppress interest rate increases is similar to the decision to impose an excess profits tax. Of course, the tax increases revenues, but increased profits, always suspect in some quarters, are especially detestable in wartime.

Economists, with a few exceptions, have found printing money an unacceptable means of war finance. But the appropriate balance between borrowing and taxation has been more controversial. Space does not permit a detailed review of the history of thought about wartime finance, but it will be useful to mention the views of a few leading economists to establish the range of the controversy. Adam Smith argued that relying mainly on borrowing was a mistake: It hid the cost of the war from the public.

> The ordinary expense of the greater part of modern governments in time of peace being equal or nearly equal to their ordinary revenue, when war comes they are both unwilling and unable to increase their revenue in proportion to the increase of their expense. They are unwilling for fear of offending the people, who, by so great and so sudden an increase of taxes, would soon be disgusted with the war; and they are unable from not well knowing what taxes would be sufficient to produce the revenue wanted. The facility of borrowing delivers them from the embarrassment which this fear and inability would otherwise occasion. By means of borrowing they are enabled, with a very moderate increase of taxes, to raise, from year to year, money sufficient for carrying on the war, and by the practice of perpetually funding they are enabled, with the smallest possible increase of taxes, to raise annually the largest possible sum of money. (Smith et al. Smith 1976 [1776], 1080)[2]

Smith then went on to explain how long-term-bond finance encourages war by hiding the true costs of war.

> In great empires the people who live in the capital, and in the provinces remote from the scene of action, feel, many of them, scarce any inconveniency from the war; but enjoy, at their ease, the amusement of reading in the newspapers the exploits of their own fleets and armies. To them this amusement compensates the small difference between the taxes which they pay on account of the war, and those which they had been accustomed to pay in time of peace. They are commonly dissatisfied with the return of peace, which puts an end to their amusement, and to a thousand visionary hopes of conquest and national glory from a longer continuance of the war. (Smith 1976 [1776], 1080),

Smith was thinking about colonial wars. When it came to the Napoleonic wars, it would be a different story. It that case, the British government did turn in the end to the printing press to help finance the war.

John Stuart Mill argued that wartime borrowing was justified as long as interest rates did not rise (Mill 1936 [1848], 873–6). Mill did not elaborate, but I would conjecture that he was guided by the analogy between private and public finance. The analogy remains influential.

[2]By perpetually funding Smith meant issuing consols. Since the bonds would never mature, taxes only had to be raised sufficiently to pay the interest on the bonds.

The modern view of optimal war finance owes a great deal to Robert Barro (1987, 1989). Barro argued that if the war government relied entirely on taxes, there the sharp spike in taxes that would result would discourage economic activity. By borrowing, the war government smooths taxes over time, minimizing disincentive effects. Smoothing interest rates, the factor that I am stressing to explain the resort to the printing press, is the opposite side of the coin.

Theorists of war finance generally rejected money finance out of hand. But it will be useful to recall some of the problems created by inflation that lead economists to reject money finance. Inflation, most economists believe, reduces the efficiency of the price mechanism as a means of allocating resources and imposes hardships on vulnerable groups including the elderly who have accumulated savings in the form of nominal incomes, and workers who enjoy rents based, for example, on long service. Memory of the costs of inflation may erode as time from the last bout lengthens. And war parties that represent debtors may be more tolerant of inflation, but it is hard to imagine any war government welcoming inflation.

The classic statement of the costs of high inflation comes from John Maynard Keynes.

> Lenin has declared that the best way to destroy the Capitalist System is to debauch the currency. By a continuing process of inflation, governments can confiscate, secretly and unobserved, an important part of the wealth of their citizens. By this method they not only confiscate, but they confiscate arbitrarily; and, while the process impoverishes many, it actually enriches some. The sight of this arbitrary rearrangement of riches strikes not only at security, but at confidence in the equity of the existing distribution the system brings windfalls, beyond their desserts and even beyond their expectations or desires, become 'profiteers,' who are the object of the hatred of the bourgeoisie, whom the inflationism has impoverished, not less than of the proletariat. As the inflation proceeds and the real value of the currency fluctuates wildly from month to month, all permanent relations between debtors and creditors, which form the ultimate foundation of capitalism, become so utterly disordered as to be almost meaningless; and the process of wealth-getting degenerates into a gamble and a lottery. (Keynes 1920 [1919], 298)

Keynes had the hyperinflations that followed World War I in mind when he wrote these passages. Long-continued secular inflation to which markets can adjust is a different matter, and likely to create fewer costs. As we will see, the bouts of inflation during America's major wars were of the short and intense variety.

The money finance, however, also comes with benefits when compared with other means of finance that costs in the minds of policymakers may offset some of the costs of inflation and make money creation an attractive alternative to other means of finance in major wars: low administrative costs, the ability to hide the costs of war, and a surprising degree of equity. Each is worth some attention.

(1) Low Setup and Administrative Costs Major wars require sudden and substantial increases in spending. The printing press can be accessed quickly and the administrative costs of collecting resources from printing money are low when compared with conventional taxation. A tax on windows requires that administrators go from house to house and count the windows, that tax bills be written and sent to householders, that money is collected, that suits are launched against householders who refuse to pay, and so on. On the other hand, when money is printed it can simply be handed to the soldier who will do the fighting. Smugglers may evade a tariff, and householders may cover the windows in their homes, but the inflation tax is hard to evade. Anyone who uses cash will see its value fall. Borrowing, it is true, can also be accessed quickly, provided the capital market is broad and deep. If the banking system is well developed, funds can be borrowed quickly from banks.

(2) Hiding the Cost of War When taxes are raised to finance wars it is all too obvious to the public that the war is costly. It is, after all, the government itself that, typically, delivers the bad news. The public may, however, not attribute the inflation resulting from money creation to the government because the chain of reasoning linking inflation to the government and the war it is waging is relatively long. It will be the shopkeeper who is forced to deliver the bad news. So the inflation that an economist might trace to money creation might be attributed by the public to a variety of other factors: shortages of particular commodities or a rising tide of greed and profiteering. Keynes makes this point in the passage quoted above when he claims that through inflation "governments can confiscate, *secretly and unobserved* [my italics], an important part of the wealth of their citizens. Keynes enlarged on this thought in a subsequent passage.

> There is no subtler, no surer means of overturning the existing basis of society than to debauch the currency. The process engages all the hidden forces of economic law on the side of destruction, and does it in a manner which not one man in a million is able to diagnose.(Keynes 1920 [1919], 298)

Keynes' estimate of one in a million was a bit of an exaggeration; but his point, that the victims of inflation may not be able to trace the inflation back to government finance, is well taken.

The means of creating new money will affect how easily the inflation is tied by the public to war finance. In the Revolution and the nineteenth-century wars, as I will show below, the public saw new forms of currency in its pocket: Continental dollars during the Revolution, Treasury notes during the War of 1812, and greenbacks in the Civil War. The argument that the government was printing the new money and that it was the new money that was causing the inflation was far from watertight. Perhaps the government was printing money simply to replace the gold and silver held by hoarders. Or perhaps it was printing money to make sure that people had enough given that prices had risen. But the connections were clearer than when subsequent forms of monetary expansion were used. Under the National Banking Act (1863), the government sold bonds to private banks that in turn issued private bank notes. This method also provided additional revenues for the government from monetary expansion and depressed yields on government bonds. But from a public relations point of view it had the advantage that the note was issued by a private bank and not by the U.S. Treasury. Once the Federal Reserve was established money could be created without alerting the public by introducing a new form of currency.

(3) Equity? In some circumstances financing through money creation may be viewed at least in some quarters as a comparatively equitable means of finance. A war party that represents debtors, for example may, at least up to a point, welcome inflation.

Inflation is essentially a tax on cash balances. The real value of cash goes down by the amount of resources acquired by the government. A tax on cash may not seem fair compared with a well-administered progressive tax on income or wealth. But compared with, say, a tax on food or fuel, a tax on cash may well be broader based and more equitable. During the Revolution Benjamin Franklin, perhaps in an over-zealous effort to defend American financial practices, recognized that inflation was a tax, and argued that since people held cash in proportion to their wealth, it was a comparatively fair form of taxation. At least in our day, of course, cash is a declining proportion of wealth, and the inflation tax is regressive.[3]

Up to this point I have been describing war finance as if decisions about the use to be made of each means of finance were made seriatim. In fact, it was an ongoing balancing act. Decisions about how much use to make of a particular form of finance were made both on the basis of previous decisions and on the basis of expectations about the alternatives that would be available. The basic principle is that the government will push each means of finance until the marginal cost in political support from raising one more dollar through a given means of finance equals the marginal cost in political support of raising another dollar from any alternative.[4] Of course, this is only the beginning of a long story. There are different forms of borrowing, different forms of money creation, and different forms of taxation each of which have different political consequences. Nevertheless, as I will try to show below, a simplification of this model which says simply that governments are afraid to let nominal interest rates increase, and when they see that prospect they turn to the printing press, goes a long way toward explaining the American experience.

So far I have assumed the traditional framework in which money creation is one of three means of financing government expenditures: taxes, borrowing from the public, and printing money. This description can be made a tautology by defining "finance" appropriately, but this tautology should not blind us to the many other

[3]More formally, money creation can be regarded as part tax and part borrowing by the government at zero interest. Start with a quantity equation (1) $M = kPy$ where M is the stock of money, k is the proportion of income held as money, P is the price level, and y is real income. Differentiating both sides and rearranging terms produces (2) $\Delta M/P = (M/P)\,(\Delta P/P) + (M/P)\,(\Delta k/k + \Delta y/y)$. The term on the left-hand side is the government's revenues from inflation. The first term on the right is the inflation tax: the product of the rate of inflation (the tax rate) and the real stock of money (the tax base). The second term on the left is the amount of money the economy will absorb before inflation sets in because more is demanded for transactions or precautionary purposes.

[4]The discussion here abstracts from the decision about the amount of resources to devote to the war effort. In reality, however, that decision will also be part of the balancing act. The war government may decide that cutting the amount of resources devoted to the war is less costly than increasing one of the means of finance.

ways governments have for getting hold of the resources they need to fight wars. Soldiers can be paid in cash to join; but they can also be given land grants, which were an important means of enlisting soldiers before the Civil War. And, of course, conscription has been used in most of America's major wars. If conscripts are paid less than a competitive wage, the difference between a competitive wage and what they are paid is an implicit tax. Still another possibility is to rely on "the kindness of strangers" as was done in the Revolution and the First Gulf War. And it may be possible to reduce current spending on civilian activities, for example by deferring maintenance on existing infrastructure. A secondary purpose of the chapter is to draw attention to the role of these non-standard means of getting hold of resources have played in "financing" America's wars.

In what follows I will argue that money creation was mainly the result of the fear that nominal interest rates would rise if borrowing from the public was increased – the distinction between real and nominal interest rates does not seem to have been a part of public decision making until recently – and that a simple version of the quantity theory of money goes a long way toward explaining the resulting inflation. Inflation, in a simple version of the quantity theory (with velocity constant) occurs at the same rate as the rate of growth of money per unit of output. Inflation, in other words, is simply "too much money chasing too few goods." It should be kept in mind, however, that the increases in the stock of money in some of these wars was extraordinary and dominated other factors affecting the rate of inflation. In peacetime, when money growth is only one of many factors, the effect of money on prices is not likely to show up in as clear a fashion as in these wartime experiences.

From the Revolution to the Civil War

The Revolution

The American Revolutionaries were taking on the richest and most powerful nation on earth. The Americans, moreover, lacked a financial system that could provide large amounts of capital in short order. There was no market for government securities and no commercial banks. Some funds were raised by domestic borrowing. Robert Morris famously lent substantial sums to the Revolutionary government, as did his assistant Haym Solomon. Toward the end of the war, the government took to paying soldiers and suppliers with warrants, obligations to pay after the war was over, a practice that sometimes was more commandeering than borrowing. Foreign loans (and subsidies) from Britain's continental rivals – the Netherlands, Spain, and most importantly France – were also a major source of finance, providing desperately needed hard currency, in the form of loans and subsidies. The French also provided aid directly in the form of military support. But all these together could not cover the costs of the war. And the foreign monies were not forthcoming in substantial amounts until the Revolutionaries could show themselves to be a substantial threat to Britain on the battlefield.

Raising sufficient revenues through taxation was also a daunting task. The Continental Congress did not have the authority to tax: It had to request revenues from the states. Taxation, moreover, was not popular to begin with – "no taxation without representation" – and many citizens, including many of the richest, were Tories who would be unlikely to cooperate willingly with the authorities. The Continental Congress therefore took to the printing press. Ease of access, low administrative costs, and the difficulty of evading the inflation tax carried the day. Overall, Farley Grubb (2011a, 272) estimated that 77 % of the expenditures of the Continental Congress were financed by issuing Continental dollars. By year: 100 % in 1775 and 99 % in 1776, when no alternatives were feasible, then 59 % in 1777, 75 % in 1778, and 77 % in 1779.

Initially, the Continental dollar, like many colonial currency issues (often to finance wars), as Grubb has emphasized, was both a zero coupon bond and currency.[5] The first issue, which was authorized within a week of Bunker Hill, bore no explicit interest but was set to be redeemed between 1779 and 1782. If the notes had been regarded simply as investments, Continental notes for $100 dollars at face value would have traded, at a safe prewar rate of say 6 %, for between $65 and $78 in specie (Grubb 2011b, 14). This prewar norm was also the rate used for coupons on government loans during the war. But some of the notes were issued in small denominations of one to eight dollars. The point clearly was that they would also circulate as money.[6] The point is sometimes made that the distinction between bond-financed deficits and money-financed deficits disappears in wartime. Here the distinction disappeared because the same instrument was part bond and part money; in later wars the distinction would disappear because money was created to fix the price of bonds.

Table 1 gives an idea of the amount of money created during the Revolution by the Continental Congress and the corresponding increase in prices. It shows cumulative issues of the Continental dollar and two measures of the price level. The amount of Continental dollars has been converted to a specie (gold and silver coin) basis using an estimate of the specie price of Continental dollars based on commodity prices in Philadelphia. This price declined dramatically during the war, giving rise to the old expression "not worth a Continental," once a common description of something that was of little value.[7]

[5] See Calomiris (1988) for a related but somewhat different interpretation of the Continental dollar.

[6] Hence, the Constitution uses the now mysterious term "bill of credit, rather than paper money or cash."

[7] Alternatively, one could think of the depreciation of the Continental dollar in specie as an increase in the Continental dollar price level.

Table 1 Money and prices during the Revolutionary War

Year	Continental dollars in circulation at face value (millions)	Specie and an estimate of the specie value of the Continental dollars (millions)	The consumer price index in specie prices	Wholesale prices
1775	6.0	12.5	100	100
1776	24.9	22.9	114	115
1777	37.9	16.1	139	164
1778	101.4	18.3	180	187
1779	200.0	17.7	160	301
1780	200.0	11.5	179	300
1781	200.0	6.5	145	288

Sources: Specie: Friedman and Schwartz (1970 table 13). I assumed that the estimate for 1775 remained in the economy and influenced spending decisions even though it may at times have been hoarded and therefore counted it as part of the money supply. Continental dollars: Grubb (2011b, table 2, p. 15). I used the price of Continental dollars that Grubb derived from a Philadelphia price index and applied it to all the outstanding issues. In fact, at times different issues circulated at different discounts because of differences in redemption features. An estimate of the total money stock would include state issues of notes and counterfeits. Consumer price index: www.measuringworth.com. Wholesale Prices: Historical Statistics (2006, series Cc113)

While the amount of Continental dollars outstanding shown in Table 1 is probably about right, the monetary situation as a whole was extremely complex. Working out the exact amount of paper money issued by the Continental Congress is a daunting task (Grubb 2008) and what became of the specie circulating before the war is a matter of conjecture. The Continental currency, as noted above, was not a simple fiat money. Rather it consisted of bonds with complex redemption schedules. Professor Farley Grubb (2011b) has worked out the implications of these redemption policies for the prices of the notes. To add to the complexity most of the states issued their own paper monies. Ratchford (1941, p. 34) estimates that the states issued an amount on the same order as the Continental dollar. There were also substantial amounts of counterfeits in circulation, including counterfeits produced by the British. The British were good at counterfeits – they had access to better facilities for printing notes than did the Americans – and printing notes undermined the American economy even as it minimized British expenditures of hard currency. Another issue concerns the treatment of specie. It is estimated that on the eve of the Revolution between 4 and 9 million dollars' worth of specie circulated in the colonies. Much of this was probably hoarded during the war as people tried to spend their fast-depreciating paper money. But it is possible that people regarded hoarded specie as part of their liquid assets, and that these hoards continued to influence spending decisions.

Figure 1 shows the percentage increases in most of the paper currencies that circulated in the United States and the percentage increases in prices when expressed in the corresponding currency. The rates of inflation on an annual or even monthly basis implied by these numbers were extremely high. The Continental dollar depreciated at a rate of about 5 % per month. The Virginia dollar depreciated 10.7 % per month. In the last few months of the Virginia inflation it appears to have come

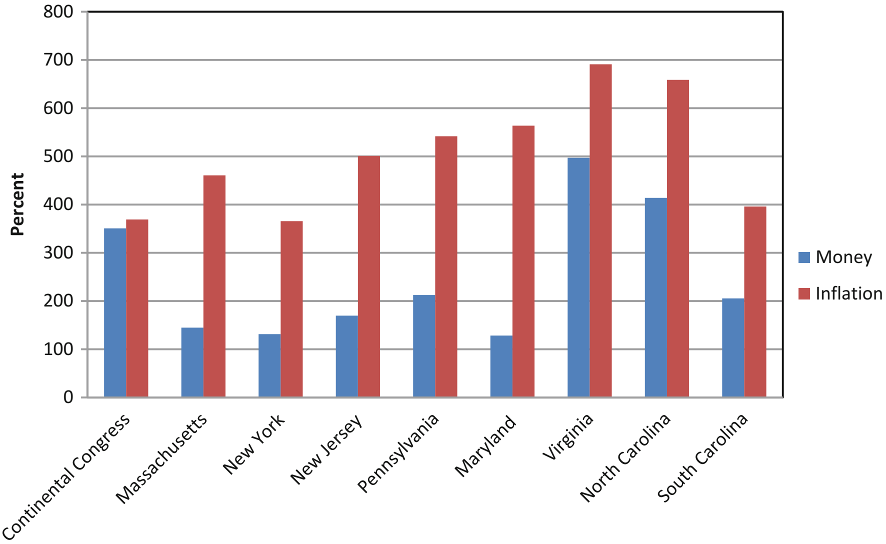


Fig. 1 Money and inflation during the Revolutionary War, percent increase from the start of the war to the price peak. Notes and Sources. Massachusetts includes New Hampshire and Rhode Island. New York includes Connecticut, and Pennsylvania includes Delaware. I assumed that the state paper monies were initially issued at par. Assuming that the state issues were initially issued at a 33.3 % discount does not alter the overall picture very much. The amount of paper money issued by the Continental Congress is from Grubb (2011b, table 2, p. 15). The amounts issued by the states are from Ratchford (1941, p. 34). Prices are from McCusker (2001, Tables C-1 and C-2, pp. 76–79)

close to Phillip Cagan's definition of hyperinflation, 50 % or more per month, which Cagan (1956) used while studying twentieth-century hyperinflations.[8]

Figure 1 shows that the depreciation of the Continental dollar was similar to the increase in the amount of Continental dollars outstanding; but the depreciation of the state currencies outran the amounts issued. I do not have a ready explanation for this difference or the differences in the experiences of the individual states. Perhaps the somewhat better performance of the national currency can be explained by differences in the redemption policies, or simply by the greater credibility of the national government. Credibility of future redemption might also explain some of the differences among the states. Virginia, for example, had very large claims to western land that might have provided a basis for redemption of the paper money, while Maryland had none. All this, of course, has to be considered highly speculative.

[8]Cagan, however, added that the 50 % rate had to persist for a year.

While paper money was the main means of finance, the Revolutionary government found other ways of acquiring the resources in the aggregate may have been equally important. The government made a start on building a navy, but the more important weapons for attacking British shipping were the New England privateers. It has been estimated that there were on the order of 10,000 people employed in building and manning privateers. After the land war moved south, privateering was New England's main contribution to the war effort. The central government did not offer cash or securities to the privateers. Instead, it promised them a share of the spoils of war: The letter of marque authorized American ships to capture British ships and offered them half the profits derived from selling the goods they took. It was an attractive offer given the high domestic prices of imported goods.

The states and the Continental Congress followed a similar strategy for recruiting soldiers for the army. Volunteers were offered land grants, to be filled with western lands that would be opened for settlement after the war. The British had restricted Western settlement in the late colonial period, but these restrictions would be removed once the war was won. Land grants were a colonial tradition, and men volunteering for the Revolution would have thought that there was a good chance that they would be rewarded in this fashion. The earliest explicit offer of land came in 1776, but ironically it was an offer of 50 acres to British soldiers who deserted (Gates 1968, 251). To make sure that the offer reached the Hessians, it was printed in German on the backs of tobacco wrappers (Hibbard 1965 [1924], 118).

Both the Continental Congress and the states made land grants. The grants made by the Continental Congress went to veterans of the Continental Army; the grants made by the states went to veterans of their militias and as supplemental grants to their citizens who served in the Continental Army. The Southern states were especially generous. Virginia, the most populist state, and the one with the most western land claims, increased its bounties during the course of the war as it became harder to enlist men. Toward the end of the war, ordinary soldiers and sailors were promised 300 acres and a slave to work the land. North Carolina made the largest grants. Under a law of 1780 the North Carolina scale started with 640 acres for privates and 1000 acres for noncommissioned officers.

The rewards to officers, in many cases, were far more generous than the rewards for enlisted men. The Continental Congress gave brigadier generals 850 acres and major generals 1100 acres. North Carolina added 12,000 acres for both brigadier and major generals. It rewarded Major General Nathaniel Greene, the hero of the Revolution in the South, with a "little dukedom," as one Tennessee legislator put it, of 25,000 acres (Gates 1968, 252–53). Virginia added an additional 10,000 acres for brigadier generals and 15,000 for major generals. Virginia granted Major General Charles Lee16,875 acres, Brigadier General Daniel Morgan 23,328 acres, and Major General Horatio Gates 31,000 acres (Bockstruck 1996, 196, 308, 376).[9] Washington undoubtedly could have had an enormous amount of land. He refused, however, to accept any land for his service during the Revolution, even though his personal

[9]The island of Manhattan is 13,000 acres.

economy depended on the extensive lands awarded to him for his service in the French and Indian War (Flexner 1984, 53–54, 194). The Northern states made smaller distinctions among ranks. Whereas Virginia gave privates 300 acres and major generals 15,000, a ratio of 50:1; New York gave privates 500 acres and major generals 5500, a ratio of 11:1 (Gates 1968, 252). The states with large claims on western lands were, of course, in a position to reward their veterans generously. The expectation that those claims might have to be surrendered to the central government simply added another incentive to reward their own citizens generously since it was likely that those claims would be honored by any national government. Although the veterans of the Revolution were promised a great deal of land, its monetary value is difficult to estimate and may not have been very great. Some preliminary efforts to distribute federal land were made in the mid-1780s, but it was not until 1796 that a military tract in Ohio was made available where small holders could exercise their warrants (Gates 1968, 259).

War of 1812

Albert Gallatin, a much-admired Secretary of the Treasury, who would serve during the first part of the War of 1812, maintained that wars should be financed mainly by borrowing with taxes raised only enough to pay the interest on the additional debt, the traditional J.S. Mill view. Gallatin's main concern seems to have been that America's trade would be disrupted in a war with Britain, and her economy could not easily absorb the further effects of high taxation (Dewey 1931, 128–29). Perhaps a connection could be made here with Barro's concern with the disincentive effects of high tax rates.

War was declared in June 1812. Immediately following the declaration Congress authorized an issue of $5 million in treasury notes (more below) and a doubling of tariff rates. But with America's foreign trade disrupted, higher tariff rates would not translate into sufficient revenues to fund payments on the debt. This was recognized from the start, but getting higher internal taxes through Congress was hard. In February 1813 Congress authorized a loan of $16 million, an amount exceeding previous authorizations and intended to finance the war. But this issue could not be sold at par, the usual prewar condition, and lower bids had to be accepted. The disappointment fostered by the need to increase yields on this loan put more pressure on Congress to increase tax revenues and to issue interest-bearing notes that could circulate as money. At a special session of Congress in the summer of 1813, a range of new taxes were levied: a direct tax (property tax) and taxes on carriages, refined sugar, stamp and auction taxes, and a variety of taxes on alcohol. In September of 1814 efforts to raise revenues along these lines were strengthened. Altogether, however, the taxes raised during the war were insufficient to maintain the credibility of the debt.

Sometimes it is claimed that the War of 1812 was financed entirely by "debt," thus making for a complete reversal from the Revolutionary War, which was financed, as we have seen, mostly by printing money. The Bureau of the Public Debt's website, for example, tells visitors that "The War of 1812 was financed mainly through the use of borrowed funds. Total public debt increased from $45.2 million on January 1, 1812, to $119.2 million as of September 30, 1815." (http://www.publicdebt.treas.gov/). But much of the debt was in small denominations intended to circulate from hand to hand as money, or to serve as interest-bearing bank reserves.[10]

As far as I am aware no one has published estimates of the stock of money during the War of 1812. The fragmentary estimates of some of the components of the stock of money to be found in the literature were assembled by Milton Friedman and Anna Schwartz (1970, Table 14). They wisely refrained from combining those estimates. Any attempt to do so faces a number of problems. For one thing, the nature of the money supply changed dramatically during the war. The Treasury Notes were one example. Were they money? Bonds? Or are they best regarded as a mixture of the two? And in August 1814, when the British captured Washington, banks outside New England suspended specie payments and state bank notes went to heavy discounts in terms of specie.[11] I have decided, however, not to follow the wise precedent set by Friedman and Schwartz, and instead put together my own estimates of the stock of money. At best these estimates are useful for painting a broad-brush picture and are subject to a wide margin of error. Nevertheless, they show (Table 2) that a substantial increase in the stock of money was accompanied by a substantial increase in prices.

Initially, veterans of the War of 1812 who served in the regular army received bounties of 160 acres, an amount that was raised to 320 acres toward the end of the war (Hickey 1989, 243–44). About 29,000 of the 60,000 regulars took them up. There were, however, several problems with the grants from the point of view of the veterans. First, the grants were located in a few restricted western areas, in part because it was hoped that the veterans would be a buffer against Native Americans. This was also true of the Revolutionary warrants, but many of the Revolutionary grants were in highly productive farmland, such as those in South Central Ohio. One area in Michigan reserved for veterans of the War of 1812 was replaced when it was found to be too swampy, although some of the other areas reserved for the veterans of 1812 did contain good farmland. Veterans who had served in state militias were not entitled to grants. State governments rewarded many veterans who had served in militias during the Revolution. However, this was no longer possible because the states had ceded their western lands to the federal government.

[10]The interest rate on these notes was set, typically, at 7.3 %. This works out to be 2 cents per day on $100, facilitating circulation from hand to hand.

[11]Presciently, Adam Smith warned in *The Wealth of Nations* about the dangers of depending on bank notes because of the risk of an enemy capturing the nation's capital.

Table 2 Money, prices, and interest rates in the War of 1812

	Government currency Issues (millions)	Total stock of money (millions)	Cost of living (specie .prices)	GDP deflator (specie .prices)	Long-term interest rate, government bonds
1811	0	56.42	100	100	7.18
1812	2.84	73.51	101	99	7.39
1813	4.91	86.85	122	115	7.61
1814	10.65	100.49	134	123	9.22
1815	14.13	130.01	117	104	8.81

Only scattered and sometimes contradictory figures are available for privately issued bank notes and deposits, so there is a wide margin of error around the estimates in the second data column. The total stock of money includes private bank notes and deposits, specie, and the government issues

Sources: Money: Friedman and Schwartz (1970, Table 13) Cost of living, GDP deflator, and long-term interest rate; www.measuringworth.com

The Mexican War

The Mexican War (1846–1848) was easily financed by borrowing; there was no need to raise taxes to assure bondholders of future interest and debt repayment, or to resort to the printing press to keep nominal interest rates down The federal budget had been in surplus before the war, so bond buyers were not going to worry about the ability of the government to meet future interest and principal commitments. All of the debt issued during the war was issued at par or at a premium. In comparing the financing of the Mexican War with the War of 1812, Dewey (1931, 255) concluded that "the ease of the treasury was due not so much to a wise intelligence as to the great increase in the wealth of the country and to the advance in government credit." Again, focusing on finance should not obscure the role of land grants which were also important in filling the ranks.

The Civil War

During the Civil War the United States created two new forms of money, the greenback and the national bank note. Although there were a variety of motives for creating them, both were created with an eye to monetizing part of the interest-bearing debt and hence keeping nominal interest rates within "normal" bounds.

(1) The Greenbacks The famous greenbacks were first issued as part of a major funding bill passed in February 1862. The war had begun in April 1861 with expectations on both sides that it would soon be over. The First Battle of Bull Run was fought in July, and by February 1862 it was clear that a long and bloody war lay ahead. The funding bill called for the issue of $500,000 in 5–20s: 6 % bonds with interest payable in gold, callable after 5 years maturing at 20. The equivalent today, as a share of GDP, would be $1.64 trillion.[12] The law also provided for $150,000 in notes: the famous greenbacks.[13] They were legal tenders: they could be used to pay taxes or settle private debts denominated in dollars.[14] A key provision of the legal tender act was that the notes were convertible into the 5–20s, so that the act created the currency with which the 5–20s could be purchased.

Secretary of the Treasury Salmon Chase does not seem to have been very enthusiastic about the greenbacks. Indeed, in 1869 as a member of the Supreme Court he would rule that the legal tender clause was unconstitutional. The idea for the greenbacks had originated in the Congress. In the Senate John Sherman argued that if the greenbacks were not issued, money for the war effort could not be had except at ruinous interest rates because gold and silver had ceased to circulate. In the House, Thaddeus Stevens was equally enthusiastic. As a result of issuing the greenbacks, Stevens argued, the $500 million in 5–20s would be subscribed before the government could use the funds. In other words, it was clearly understood that the greenbacks were a vehicle for monetizing the debt and preventing an increase in nominal rates.

There would be two further authorizations of greenbacks in July 1862 and July 1863: $450 million in all, with about $430 million issued.

The $500 million in 5–20s authorized in February 1862, did not sell well initially because Chase insisted on selling them at par even though the legislation authorizing them had authorized sale at market price. Sales, however, did pick up for a number of reasons, including the depreciation of the greenback, which made the 5–20s, which promised interest in gold, more attractive. The 5–20s could be exchanged at par for greenbacks, so essentially the government was printing money to buy bonds; economically the same as the Federal Reserve open market operations undertaken during World War II to be discussed below. Eventually, the right to convert greenbacks into interest-bearing gold bonds was terminated, so the greenbacks became a pure fiat money.

The reputation of the greenbacks was mixed. The greenbacks were popular in Congress. And in parts of the country "Lincoln Green" was seen as a superior form of money. This was especially true in the West where some currency issued by state banks had been backed by Southern bonds and so had lost most of its value. On the other hand, it was widely recognized that the greenback was a source of inflation. This was more obvious than it would be in subsequent wars because there was an active market for gold dollars and one could see the increase in the greenback price of gold from day to day.

---

[12] Using the calculator at www.measuringworth.com

[13] They were known as greenbacks, of course, for their color which became traditional for American currency, and perhaps also because they were "backed" by green ink rather than by gold.

[14] Except for tariffs, which had to be paid in specie. Allowing tariffs to be paid in greenbacks would in effect have lowered the tariff rate, something anathema to the Republicans.

(2) The National Banking System Chase's preferred method for dealing with the currency was a national banking system. He had mentioned this idea as early as his report to Congress in December 1861. The basic idea was a nationalized version of the "free banking law" that had been adopted before the war in a number of states including Chase's home state of Ohio. The key provision, for our purposes, was that paper money would be issued by private banks, but it would be backed by government bonds. Under the state-level free banking laws, states chose different lists of eligible bonds, including of course bonds of the state in which the bank was located. Under Chase's national banking act, the federal government would charter banks, and their notes would be backed by federal bonds. Chase offered a number of reasons for supporting a national banking system. For one thing it would provide a more uniform and safer currency than the myriad of state banking systems that had existed before the war. At the same time since these notes would be issued by private banks and redeemable in gold or silver (after resumption of specie payments) there would be less danger of inflationary issues than if the power to issue paper money were permanently left in the hands of the Congress. Clearly, however, monetizing part of the federal debt and thus keeping nominal interest rates low was an important gain expected from the National Banking Act.

The consequence of the increase in the stock of money during the war was inflation. This can be seen in Table 3, which shows the issues of the greenbacks, estimates of the total stock of money, two measures of the price level, and the interest rate on long-term federal government bonds during the war years. Evidently, the stock of money rose rapidly in the North during the Civil War.[15] The cost of living rose by a factor of 1.9 and the GDP deflator by a factor of 1.7.

Inflation was disruptive and undermined support for the war. For one thing creditors suffered unexpected losses, thus harming an important constituent of the

Table 3 Money, prices, and interest rates in the Civil War

	Amount of greenback dollars outstanding (millions)	Total stock of money (millions)	Cost of living	GDP deflator	Long-term interest rate, government bonds
1860	0	554	100	100	5.57
1861	0	558	106	104	6.45
1862	149	729	121	116	4.92
1863	411	973	151	143	4.37
1864	471	1397	189	175	4.83
1865	456	1445	196	178	5.54
1866	428	1315	191	169	5.52

The monetary totals are for June. The total stock of money includes greenbacks, specie, interest-bearing federal currency, bank notes, and bank deposits
Sources: Money, *Historical Statistics* (2006, series Cj28). Cost of living, GDP deflator, and long-term interest rate; www.measuringworth.com

[15]The inflation was more severe in the South, which relied more heavily on the printing press.

Republican party.[16] Wages did not keep pace with the inflation although there is a debate about whether this was a purely monetary phenomenon, or whether is it was also the result of real disturbances such as the cutoff of supplies of cotton. One of the groups badly hurt by wartime inflation was the soldiers. Wages of soldiers remained at $13 per month from the start of the war until May 1864 when they were increased to $16 per month, an increase that was too small to catch up with the current price level, let alone make up for past losses. The problem of rigid wages for soldiers combined with rapid inflation would also prove to be a problem in World War I.

Emphasis on money creation should not distract attention from the efforts the Lincoln government made to fund and market its debt. The Republicans were serious about taxes. In addition to raising the tariff rate, the Republicans imposed a series of excise taxes. The taxes they imposed on alcohol and tobacco are still with us. Perhaps even more remarkable, they imposed progressive taxes on corporate and personal incomes and inheritances. To help market the debt, Chase turned to the investment banker Jay Cooke. Cooke assembled a team that fanned out across the North and sold bonds to doctors and lawyers and other middle-class investors. Constructing new means of borrowing and taxation took time. But by the end of the war, reliance on new issues of greenbacks had ended. Although the inflation was substantial, it was nothing like the inflation in the South during the Civil War, which relied almost entirely on the printing press.

Nonfinancial methods of acquiring resources were also used during the Civil War. Conscription was the main form in this class. As in the Revolution, conscripts could hire substitutes. Thus the amount paid to substitutes could be regarded as a tax on the conscripts. It simply went straight from taxpayer to soldier, bypassing the government bureaucracy. Land grants were not an important part of war economics. The Homestead Act, however, which was passed during the war, allowed veterans to count their war service toward the 5 years of residence required under the Act to obtain title to the land.

The consequences of relying or not relying on the printing press in the nineteenth century can be seen in Fig. 2 which shows the percentage increase in prices and the percentage increase in money per unit of output in five nineteenth-century wars. The price series used for Fig. 2 is a consumer price index. But the GDP deflator when available produces a similar picture. Money per unit of output is the percentage change in the stock of money divided by real GDP. To be sure, real GDP estimates from the antebellum and Civil War years are best described as conjectures. The estimates of the stock of money for the War of 1812 and the Civil War are my own. The estimates are useful for broad-brush comparisons but no more because a good deal of data is missing and because of the added complication of different kinds of money circulating at varying prices rather than at face value.

Overall, the results are consistent with the notion that inflation was generated in major wars by resort to the printing press. There were large increases in money per unit of output in the Civil War and substantial inflation. In the War of 1812

[16]But western farmers, another important constituent of the party, one imagines, were debtors.

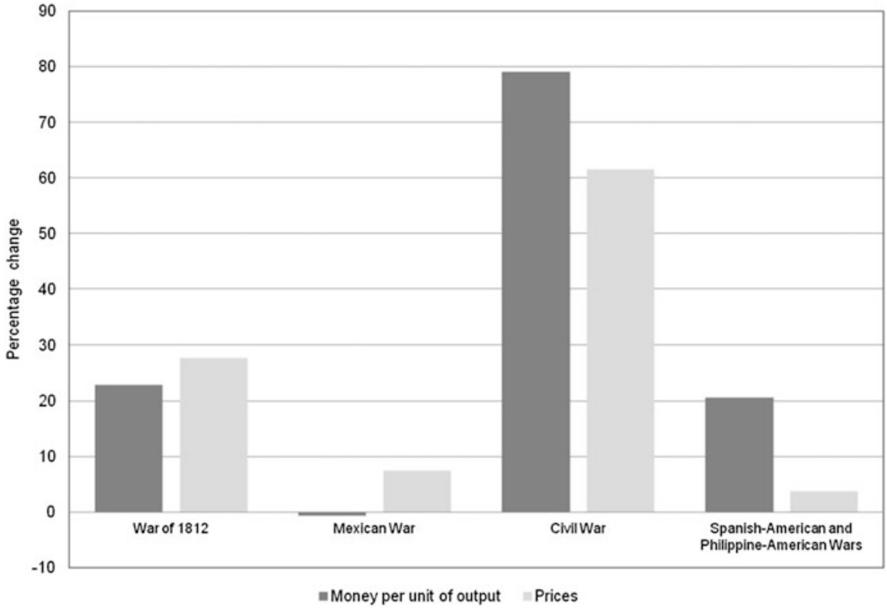


Fig. 2 Money per unit of output and prices in nineteenth-century wars. Notes and Sources. The war periods were defined as the start of the war to the price peak: War of 1812, 1812–1814; Mexican War, 1846–1847; Civil War (North), 1861–1865; Spanish-American and Philippine-American Wars, 1898–1902. The sources for money, prices (consumer price index), and output (real GDP) are given in Tables 1, 2, and 3. Percentage changes were calculated by taking the difference in natural logarithms and multiplying by 100

the increases in money per unit of output and inflation were somewhat less. In the Mexican War there was no resort to the printing press and no inflation.

The most surprising result is for the Spanish-American and Philippine-American wars. Money per unit of output did increase noticeably, but prices increased to a much smaller extent. One possible explanation is that the Spanish-American and Philippine-American Wars occurred during a period of economic slack similar to what we are experiencing today. To be more precise, the Spanish-American War began 5 years after the great panic of 1893 – just as this chapter is being written in 2013, 5 years after the panic of 2008. And as we are acutely aware there is still considerable slack in the economy. Real GDP recovered rapidly after the panic of 1893. By 1895 real GDP was higher than before the crash. But getting back to full employment was a long, drawn-out affair. Figure 3 shows unemployment beginning 3 years before the crisis of 1893 and 3 years before the crisis of 2008.

After the panic of 1893 unemployment, as shown in Fig. 3, did not get back to the pre-panic level until 1901. There is a larger drop in the unemployment rate between 1898 and 1899 than in earlier post-1893 years that may be associated with spending on the Spanish-American War. There was also the unusually large increase in the money supply shown in Fig. 2. The reason appears to be an increase in the stock of

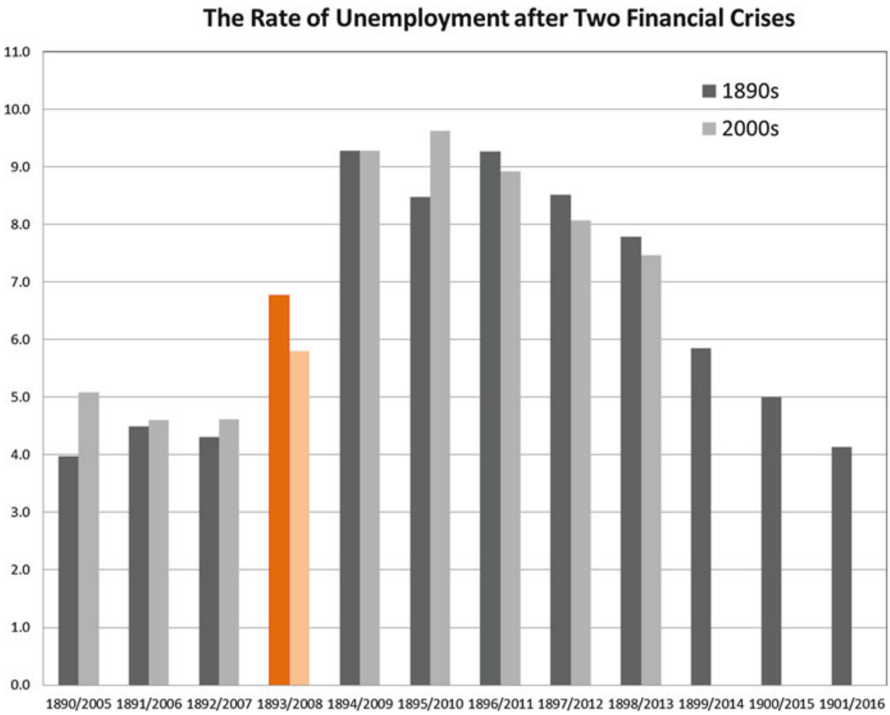


Fig. 3 Rate of unemployment after two financial crises (Sources: Historical Statistics of the United States, Millennial Edition; U.S. Bureau of Labor Statistics)

monetary gold: Favorable harvests in the United States combined with poor harvests in Europe produced an export surplus and a substantial inflow of gold. Thus, in the 1890s the gold standard produced by chance the kind of monetary stimulus that the Federal Reserve is now attempting to produce deliberately. If the increase in the stock of money did not produce an increase in commodity prices, what did it do? It may account for the drop in the unemployment rate. It may also have fed into the stock market, which enjoyed a bull market – until the stock market panic of 1901. Clearly, unemployment after the 2008 panic has followed a path that is eerily similar to the path followed by unemployment after the panic of 1893. Hopefully we will do better than we did in the 1890s, but if we continue to follow the earlier path we will not get back to the pre-panic level of unemployment until 2016.

The World Wars

The establishment of the Federal Reserve in 1913 substantially changed the mechanics of wartime money creation. In the Revolution, the War of 1812, and the Civil War the phrase "printing money" was literally true; after 1913 it became a metaphor. The Continental dollar, the Treasury notes from the War of 1812, and the greenbacks from the Civil War were all freshly off the press. They included some bond-like features, but there was no mistaking the fact that they were intended to circulate from hand to hand as money. Reliance on new types of money made the chain of reasoning connecting the inflation to the monetary policies of the war governments fairly straightforward. Many people undoubtedly blamed rising prices simply on the rising greed of the shopkeepers, or the disruptions of war. But it was not hard to see that specie had disappeared, that the government was printing the new stuff that replaced the specie, and perhaps that the government was more than replacing the specie.

With the Federal Reserve in place, however, the process became more roundabout. The money supply consisted mainly of Federal Reserve notes and bank deposits, although gold, subsidiary coins, and some additional historical paper remnants continued to circulate. The amount of Federal Reserve notes and deposits increased, but there was no new currency to signal the change in monetary policy. The basic stories of monetary policy in World War I and World War II were similar. But there were important differences in the administrative details.

World War I

America declared war on April 6, 1917. Eighteen days later Congress authorized borrowing of $5 billion. This would be 1.52 trillion as a share of today's (2012) GDP, about the same effort that the United States made after the First Battle of Bull Run when an understanding of the magnitude of the Civil War was realized. Evidently, Congress well understood the size of the task that lay ahead. This was the total authorization; it was assumed that the Secretary of the Treasury would space out the actual bond sales as the war progressed. One of the stipulations was that the coupon not exceed 3.5 %. Clearly, Congress saw a low rate of interest, reflecting prewar norms, as important. As can be seen in the last column of Table 4 this coupon would have meant selling the bonds below par if typical prewar rates prevailed.

A number of measures were taken to increase the market for the bonds. In October 1917, after considerable wrangling, taxes were raised. Many of the new or increased taxes were excises. Alcohol, tobacco, jewelry, cameras, cosmetics, chewing gum, and many others came in for new or increased taxes. Income taxes, now possible because of the Sixteenth Amendment, were raised. The highest rate was 67 % on incomes over $2,000,000, about $122,000,000 today (2012).[17] There was also a graduated tax on inheritances and an excess profits tax. Although the actual policy followed with respect to taxes was similar to that followed in the Civil War and was consistent with the classical idea of war finance that borrowing was OK

[17]Using the unskilled wage as the inflator. Using a price index would yield a lower figure of about $35,000,000. Data from www.measuringworth.com

Table 4 Money, prices, and interest rates in World War I

	High-powered money (billions)	Total stock of money, M2 (billions)	Cost of living	GDP deflator	NNP deflator	Long-term interest rate, government bonds
1913	3.417	15.73	100	100	100	4.74
1914	3.532	16.39	101	101	101	4.70
1915	3.669	17.59	101	104	105	4.78
1916	4.178	20.85	111	117	118	4.69
1917	5.096	24.37	134	145	146	5.05
1918	6.190	26.73	157	169	168	5.52
1919	6.770	31.01	180	173	170	5.49
1920	7.368	34.08	209	197	194	6.12
1921	6.679	32.85	186	168	166	5.97

High-powered money (monetary base) is currency in 1913 and currency plus bank deposits at the Federal Reserve from 1914 on. The total stock of money includes greenbacks, specie, national bank notes, federal reserve notes, and bank deposits held by the public in commercial banks
Sources: High-powered money, money, and NNP deflator: Friedman and Schwartz (1982, Table 4.8). Consumer price index, GDP deflator, and long-term interest rate: www.measuringworth.com

as long as interest rates did not increase, I should note that the ideas of economists were beginning to change. Some economists were advocating more reliance on taxation. The primary argument emphasized equity. If you rely on bond finance, then young men are taxed twice, once when they are conscripted into the army, and once again when they come home and pay higher taxes to pay interest and principal on the bonds. Better to tax the stay-at-home fat cats than sell them bonds.

The Treasury also made efforts to encourage people to buy bonds through a national campaign based on patriotism. For one thing, giant rallies were held in which celebrities, including Hollywood stars, urged people to support the war effort by buying the bonds.[18] Even the name of the bonds reflected the Treasury's new emphasis on salesmanship. These were "Liberty bonds" not the prosaic 5–20s of the Civil War.

But the Treasury also relied on the Federal Reserve to monetize part of the debt. For short-term financing, which was needed before higher taxes could be realized or between sales of bonds on the open market, the Treasury sold short-term certificates of deposit directly to the Federal Reserve. Reliance on the Federal Reserve, however, went much further. The Federal Reserve set a discount rate on loans secured by government bonds that was below the yield on the government bonds. Thus, banks were encouraged to loan money to the public with which the public could buy Liberty bonds. Although the institutional details were different the effect was similar to the issue of greenbacks with the right of conversion into 5–20s. The effect of these operations can be seen in the first column of data in Table 4,

[18] Kang and Rockoff (2006) discuss in some detail the efforts to market the bonds.

which shows the amount of high-powered money (currency plus bank deposits at the Federal Reserve) during the era of World War I.

The amount of high-powered money created by the Federal Reserve, however, does not tell the whole story. Since banks only keep a fraction of the reserves they obtain, there is a multiple expansion of deposits and loans. During the war years one can assume that many, perhaps most, of these loans went for the purchase of government bonds or freed up other funds that could be used to purchase government debt. The second column of data in Table 4 shows the increase in the total money supply. The percentage increase in the money supply from 1916 to 1920 was about the same as the percentage increase in high-powered money. But the amount of the increase was much larger for money. Since many government bonds ended up on the balance sheets of individuals rather than financial institutions, it is hard to know exactly how much of the purchases of the war debt was financed by money creation. Friedman and Schwartz (1963, 221) argued that virtually all of the increase in the money supply directly or indirectly went to finance the war effort.

The decision by the Federal Reserve to support the market for Liberty Bonds seems to have been taken for granted by both the Treasury and the Federal Reserve during the war.[19] Continuation of the policy afterwards, however, became a contentious issue. The Treasury, under Secretary Carter Glass, wanted a continuation of the low discount rate that prevailed during the war to ease Treasury refunding, that is to prevent an increase in the yields on government bonds. For a time the Federal Reserve acquiesced. The Federal Reserve did not want an open confrontation that might have led to a transfer of power to the Treasury, and the Federal Reserve was concerned that a sudden decline in government bond prices might weaken the banks because they held large amounts of bonds on their own account and loans secured by bonds. Benjamin Strong, the President of the New York Federal Reserve, appears to have been the leading advocate of a higher discount rate to stem the inflation that continued after the war (as shown by the price indexes in Table 4). Rates were finally raised in January 1920 to the then very high rate of 6 %. Inflation, gold exports, and the possibility that the United States might be driven off the gold standard tipped the balance. Even Carter Glass was persuaded that an increase was needed.

World War II

The Federal Reserve supported the bond market immediately after Pearl Harbor. A few months later, Mariner Eccles, the chair of the Federal Reserve Board, explained that the Federal Reserve had bought about $100 million of government bonds after Pearl Harbor because the market had been "very weak, very jittery" (Rockoff 2012,

[19]This paragraph is based mainly on Friedman and Schwartz (1963, 221–231).

170).[20] The question from the start was not whether the Federal Reserve would support the bond market, but rather at what price. In April 1942 the Federal Reserve announced that it would keep the interest rate on Treasury bills (generally 90 days) fixed at 3/8 of 1 %. Although there was no explicit set of rates for longer maturities, a rigid pattern emerged. The Federal Reserve, in other words, promised to set a floor under the price of government securities and buy whatever amount of bonds was necessary to prevent the price from falling below that level. The yields were low compared with yields in the nineteenth and early twentieth centuries, or in the postwar period. But they reflected what had happened in the Great Depression. A low level of economic activity and the rush to government securities for safety meant that for a decade rates on government bonds had remained at very low levels. The Federal Reserve simply promised to preserve that structure of rates during the war. From a long historical view the rates in the 1930s were abnormally low, but from the perspective of 1941, they were the “new normal.”

Deliberate purchasing of debt by the Federal Reserve was a departure from what was done in World War I. In World War I, as I discussed above, the Federal Reserve monetized bond issues mainly by lending money to banks on paper collateralized with government bonds. It bought relatively small amounts for its own account. In World War II, however, the Federal Reserve went into the market and bought government bonds. Both methods were effective. In both cases the expansion of bank reserves allowed banks to purchase government securities or to make loans to customers who in turn bought bonds, either directly with the funds lent to them, or indirectly by using other funds freed up by the loan (Table 5).

I do not have a clear answer for why the Federal Reserve’s policy changed. Friedman and Schwartz (1963, 561–2) suggest that it was “dissatisfaction with the World War I experience” that explains the change, but they do not elaborate. Another possibility is that the change reflected changes in Federal Reserve operating procedures that were already in place. When the Federal Reserve was founded in 1913 it was natural to emulate the Bank of England, which generally relied on discount policy. This was all the more necessary because the U.S. government did not have a large debt and much of the debt that it did have was used by banks to back national bank notes. So a purchase of securities might produce perverse offsetting effects on the amount of cash in circulation. The large increases in debt during World War I and the Great Depression made reliance on open market operations easy for the Federal Reserve.

Friedman and Schwartz (1963, 563) claimed that “the support program converted all securities into the equivalent of money.”[21] Here we can see the strong family resemblance between the financing of the Revolution with debt that was intended to

[20] Eccles was appointed by Roosevelt and generally supported the Keynesian (perhaps more accurately proto-Keynesian) view that the main function of monetary policy in most circumstances was to keep interest rates low.

[21] Although they did not modify their estimates of the stock of money to include government debt during the period when bond prices were fixed by the Federal Reserve.

Table 5 Money, prices, and interest rates in World War II

	High-powered money (billions)	Total stock of money, M2 (billions)	Cost of living	GDP deflator	NNP deflator	Long-term interest rate, government bonds
1939	17.501	49.27	100	100	100	3.01
1940	21.358	55.20	101	101	101	2.84
1941	23.341	62.51	106	108	109	2.77
1942	25.427	71.16	117	117	123	2.83
1943	30.181	89.91	125	123	140	2.73
1944	35.788	106.82	127	126	150	2.72
1945	41.851	126.63	130	129	157	2.62
1946	44.241	138.73	141	144	158	2.53
1947	45.026	146.00	161	160	171	2.61
1948	46.166	148.11	173	169	182	2.82
1949	45.396	147.46	172	169	180	2.66
1950	43.642	150.81	173	170	183	2.62

High-powered money (monetary base) is currency plus bank deposits at the Federal Reserve. Money includes currency held by the public and bank deposits held by the public in commercial banks. The NNP deflator includes adjustments that Friedman and Schwartz made for price controls
Sources: High-powered money, money, and NNP deflator: Friedman and Schwartz (1982, Table 4.8). Consumer price index, GDP deflator, and long-term interest rate: www.measuringworth.com

circulate as cash, and with the financing of the Civil War with greenbacks that could be converted into government bonds on demand. The increases in high-powered money, and the stock of money held by the public about doubled during World War II (Table 5) and prices, as shown by several indexes, rose substantially.

The Postwar Era

After World War II the United States fought smaller colonial wars. For the most part, it was able to avoid funding those wars with the printing press. In both the Korean War and the Vietnam War, however, the administrative branch of the government put strong pressure on the Federal Reserve to expand the money supply and support government bond prices.

Korean War

The World War II bond support program remained in place after the war. Although it may have contributed to the postwar surge in prices, debate over it remained muted until the outbreak of the Korean War in June 1950. The Federal Reserve believed that maintaining a floor under bond prices had converted it into an engine of inflation during World War II, and it did not want to repeat the process in the

Korean War. Prices had increased by more than 50 %, and inflation in turn had led to the introduction of an annoying system of price and wage controls and rationing. The Treasury was insistent, however, that full support for securities markets be continued. The Federal Reserve made a tentative move toward independence in August 1950 when it announced that it would permit some increase in the yields on government securities. The Treasury reacted quickly by announcing that it would maintain current yields in its next refunding operation.

The conflict came to a head in March 1951. Conferences were held between the two agencies and with various members of the government. The result was the famous Treasury–Federal Reserve Accord which relieved the Federal Reserve of the obligation – or so the agreement was interpreted by the Federal Reserve – to peg the price of government securities. Even President Truman was involved. In his *Memoirs* Truman (1956, volume 2, 44–45) discusses both his involvement and his opposition to the Federal Reserve's interpretation of the Accord. Truman, as he clearly lays out, did not want to see higher interest rates because of the impact higher rates would have on public opinion.

> It did not seem appropriate to me that we should enter into a period of deficit financing on a rising money-rate pattern. I also felt strongly that in the moment of impending crisis we should not take deliberate steps that could possibly disturb public confidence in the nation's financing. . . . For that reason I invited members of the Federal Reserve Board to visit with me. I asked them to give the Treasury their full support for its financing program, just as they had done during World War II. . . .
>
> I was given assurance at this meeting that the Federal Reserve Board would support the Treasury's plans for the financing of the action in Korea. This assurance was given entirely voluntarily. At no time during the conference did I attempt to dictate to the Board or tell them what specific steps they ought to take. I explained to them the problems that faced me as Chief Executive, and when they left I firmly believed that I had their agreement to cooperate in our financing program. I was taken by surprise when subsequently they failed to support the program.

Although the Accord freed the Federal Reserve from the obligation to peg the price of government bonds, the Federal Reserve did not move to the opposite extreme and abandon the bond market to its own devices. The Federal Reserve continued to buy some bonds and exercise a moderating influence on interest rates. It was not until March 1953 that pegging was explicitly rejected.

Why did the Federal Reserve refuse to maintain peg bond prices as it had done in World War II? One factor was that inflation and price controls were still fresh in everyone's mind. In that respect the situation is reminiscent of the period after the Revolution. The memory of inflation during the Revolution, and of the problems generated by state currency issues under the Articles, led to the adoption in the Constitution of an explicit ban on state issues of paper money, and perhaps (the question is controversial) a ban on federal issues.[22]

[22] The Constitution grants the federal government the right to coin money (presumably gold and silver) and to issue notes (interest-bearing securities). It is silent on federal bills of credit (paper money).

Friedman and Schwartz (1963, 626–27), however, see the Accord as a response to the growing worldwide abandonment of Keynesian cheap money policies because in country after country the anticipated postwar problem, a return to the depressed conditions of the 1930s, was replaced by the problem of persistent inflation. In other words, Friedman and Schwartz saw the Accord as the product of an intellectual change. This seems somewhat early to me; after all, Friedman and Schwartz had not yet attacked the Keynesian orthodoxy! But it may be that attitudes were already beginning to change.

Another part of the story may be the size of the financial problem facing the government. Although China entered the war in October 1950 it remained a limited war. Substantial tax increases were put in place in the revenue act of September 1950, as a result of an overwhelming vote in the Congress: 328 to 7 in the House. These included increases in income taxes and corporate taxes, as well as "sin" and luxury taxes. All in all the enthusiasm of the Congress for taxes was most reminiscent of the tax bill passed shortly after the start of the Spanish-American War (Rockoff 2012, 247). It was, as Bank et al. (2008, 115–16) pointed out, a "rally round the flag" vote. Fortunately, these tax increases turned out to be enough. Federal spending increased from $40 billion in 1950 to $74 billion in 1953 – that is from about 13 % of GDP to 20 % of GDP. But federal revenues about kept pace, increasing from 13 % of GDP in 1950 to 18 % in 1953. In World War II, by way of contrast, Federal spending increased from 9 % of GDP in 1940 to 44 % in 1945. Revenues also increased, from 6 % of GDP to 20 %, but that still left a large gap, equaling almost a quarter of GDP to be filled by borrowing or printing money. Had it been necessary to increase spending to anything approaching the same extent during the Korean War, the government might well have been forced to issue more debt and Truman would have found it necessary to increase pressure on the Federal Reserve to prevent "a rising money-rate pattern." As it was, the Federal Reserve permitted an increase of the long rate from 2.62 % in 1950 to 3.20 % in 1953. Hardly a staggering increase, but sufficient to show that it had achieved a modicum of independence.

Vietnam War

There were considerable increases in money and inflation during the era of the Vietnam War, as shown in Fig. 4. War finance was a contributing factor but not the whole story.

Events moved swiftly after Lyndon Johnson became president. He moved ambitiously to create his Great Society with bold initiatives to help the poor and middle class. In January 1964 Johnson declared his War on Poverty, but in August 1964 came the Gulf of Tonkin Resolution authorizing the use of force against North Vietnam. Like Harry Truman, Johnson was determined to prevent a communist takeover of a divided country in South East Asia. The Kennedy–Johnson tax cut had been put in place in February 1964, a triumph for the "New Economics" of his Keynesian advisors. But now with rising government expenditures and rising

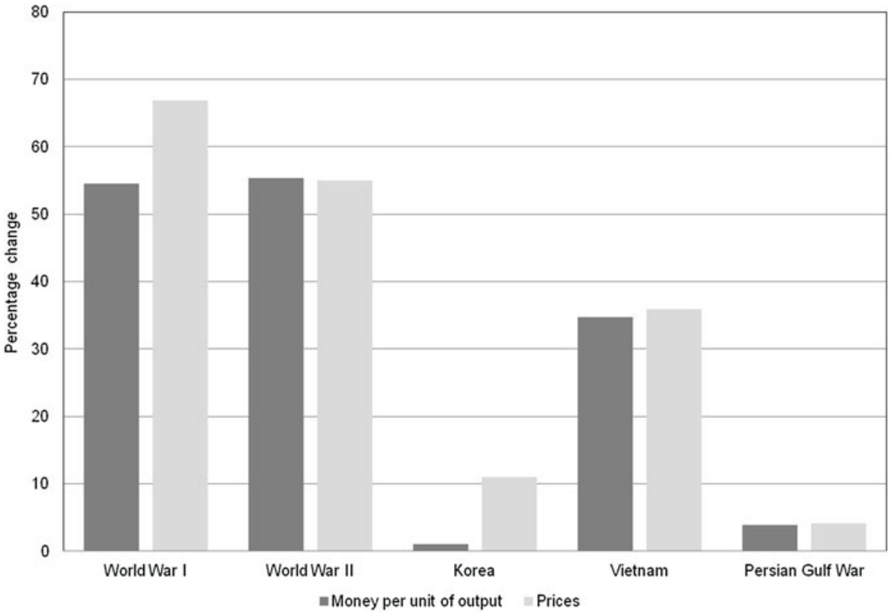


Fig. 4 Money per unit of output and prices in twentieth-century wars. Notes and Sources. The war periods were defined as the start of the war to the price peak: World War I, 1914–1920; World War II, 1939–1948; Korean War, 1950–1954; Vietnam War, 1964–1973; Persian Gulf War, 1990–1991. The sources for money, prices (consumer price index) and output (real GDP) are given in Tables 4 and 5. Percentage changes were calculated by taking the difference in natural logarithms and multiplying by 100

prices on the horizon, his economic advisors began to talk about tax increases. For a time Johnson resisted, afraid that tax increases would undermine support for both his Great Society programs and his war in Vietnam. Eventually, however, Johnson came around to the idea that a tax increase was necessary. He proposed a "surtax," an additional 10 % on personal and corporate incomes, that was enacted in June 1969.[23]

Meanwhile, the Federal Reserve was also becoming concerned about the first stirrings of the inflation that would bedevil the economy for the next decade and a half. The result was a clash between President Johnson, who wanted the Federal Reserve to maintain interest rates despite the growing deficit and rising rate of inflation, and William McChesney Martin, the chair of the Federal Reserve. Martin, although not an extreme inflation hawk, is still remembered for his description of the job: to take away the punch bowl just when the party is getting interesting. But one dramatic episode brought the conflict into sharp focus. Early in December 1965

[23]Bank et al. (2008, 136) describe the details of the tax. 10 % was an exaggeration because it applied for only part of the tax year.

the Federal Reserve announced an increase in the discount rate (the rate at which the Federal Reserve lends to member banks) from 4 to 4.5 %. Martin knew that Johnson was opposed to any increase in rates, but felt that the time to act had arrived. Johnson publicly condemned the move (*New York Times*, December 6, 1965, 1). "I regret, as do most Americans, any action that raises the cost of credit, particularly for homes, schools, hospitals and factories." In the Senate a number of prominent Democrats denounced the move and called for hearings. In the House, the legendary Congressman Wright Patman of Texas, the chair of the House Banking Committee, called on the President to demand Martin's resignation "to prevent the country from being thrown into economic crisis." Martin was summoned to Johnson's Texas ranch and there for two hours they discussed the Federal Reserve's decision (*New York Times*, December 7, 1965, 1). The press conference that followed was suffused with sweetness and light, but according to Martin's later memory, it was not all sweetness and light while the doors were closed (Meltzer 2009, volume 2, book 1, 458.) The parallels between this episode and the meeting between Truman and officials from the Federal Reserve described in the previous section are striking. In both cases the Federal Reserve was subjected to intense pressure from a populist president who was vehemently opposed to any increase in nominal rates.

As was true after the Korean War Accord, the Federal Reserve, although it had successfully asserted its independence, did not follow a sharply restrictive monetary policy. Open market operations continued to be expansionary and high-powered money, and the money supply continued to grow. The policy stance of the Federal Reserve reflected Martin's concern that policy reflect a consensus on the Federal Reserve Board, and then as now there were inflation doves as well as hawks, and Martin's understanding that the Federal Reserve was independent within (not from) the government.[24]

The inflation continued through the 1970s. And it obviously cannot all be blamed on the war. Federal spending was rising for domestic reasons as well as because of the war. And there was an important change in monetary policy. Before World War II peacetime monetary policy was aimed mainly at maintaining the gold standard. After World War II price stability remained a major determinant until the 1960s. As Keynesian economics became the dominant paradigm, mainstream economists and policymakers came to believe that the Federal Reserve could reduce unemployment through monetary expansion with an acceptable impact on inflation. As time went by it became clear that the gains in employment from monetary expansion were temporary, and trying to suppress unemployment through monetary policy produced unacceptably high rates of inflation.

The relationship between money per unit of output and inflation (consumer price index) in America's twentieth-century wars is shown in Fig. 4. As expected, the World Wars witnessed the largest increases in the stock of money relative to real output and the largest increases in prices. Perhaps the most surprising aspect of the

[24]Meltzer (2009, volume 2, book 1, 441–79 is a detailed account of Federal Reserve policy in the mid-1960s.

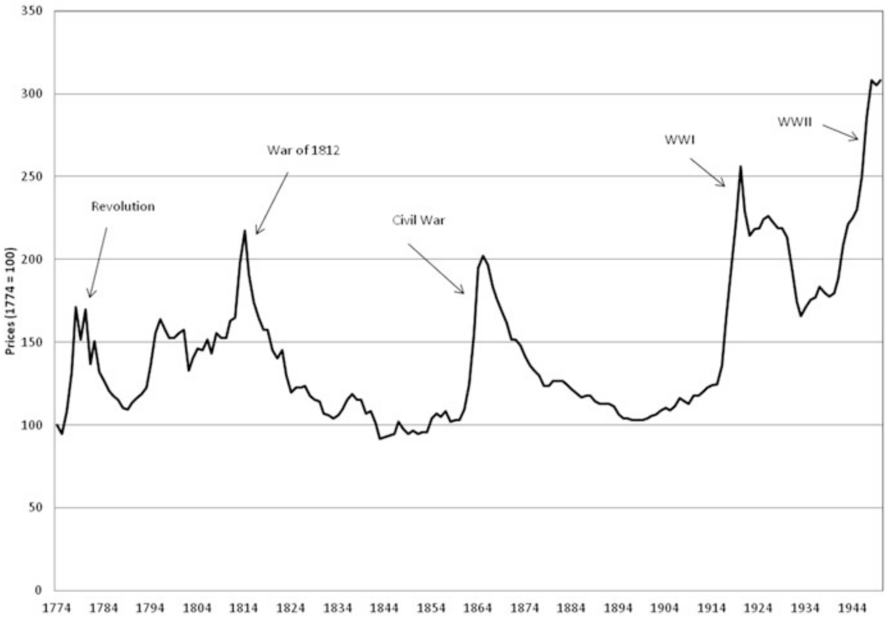


Fig. 5 Consumer prices, 1774–1950 (Source: www.measuringworth.com)

figure is that the increases of money per unit of output and prices were similar in World War I and World War II, even though World War II lasted much longer and produced far more casualties. Part of the explanation is that taxes financed more of World War II than World War I, taxes yielded more revenue in World War II, leaving less work for monetary expansion. The attack on Pearl Harbor created deep and long-lasting support for the war, making it possible for the Roosevelt administration to increase taxes without worrying about adverse political effects.

By way of summary we can look at Fig. 5 which shows the consumer price index from 1774 to 1950. During that long period inflation was almost completely a wartime phenomenon. Until World War I wars were followed by deflations that returned the price level to the prewar level. The price level in 1900 was about the same as it had been in 1774. This was the result of adherence to the gold standard and to the determination of governments to return to prewar prices of foreign currencies after the war was over. Wartime inflation, however, was not fully reversed between World War I and World War II, and after World War II we entered a new world of fiat money in which some measure of inflation became the norm. It is not too much of an exaggeration to say that after World War II what had once been the unique finances of wartime – high levels of government spending financed in part by borrowing from the public and in part by money creation – became the peacetime norm.

Conclusions

The operating assumption here has been the classical, perhaps old-fashioned idea, that inflation is caused by "too much money chasing too few goods." The evidence shown here, of course, does not prove that proposition; at best it strengthens it a bit by adding a few consistent but contestable observations to a large body of other observations with which monetary historians are familiar. In the wars there were various non-monetary factors at work, such as the disruption of foreign trade and large increases in government spending, which could also explain wartime inflation, and these factors undoubtedly did have some impact. Many years ago Milton Friedman (1952) pointed out that the inflation and increase in money per unit of output during World War I and its aftermath was about the same as in the Civil War and World War II, even though the dislocations measured by duration, casualties, government spending, and so on were far greater in World War II and the Civil War than in World War I. That set of comparisons obviously strengthens the case for assigning a primary role to money. Here we have added some examples from the pre-Civil War and post-WWII periods, but while they are consistent with money being the major determinant of the rate of inflation, they do not add decisive evidence. Inflation and money growth were less, for example, during the War of 1812 than in the North during the Civil War, but that would be expected both on monetary and non-monetary grounds.

America's wars with major powers have been financed to a significant extent by the printing press, and have produced significant inflations. Its wars against second-tier powers were not financed in this way, and did not produce inflation. Only the Vietnam War is a partial exception. Although the political economy of each war was unique, there were some common features. The natural reaction when faced with a major war was for governments to borrow the sums needed. But large-scale borrowing raised the prospect of substantial increases in interest rates. For a variety of reasons war governments were loath to see interest rates rise above prewar norms. For one thing, higher rates would be a signal to the public and to friends and foes abroad that the government's decision to wage war was undermining the economy. Increasing taxes at least to a level that promised to be sufficient to pay interest and principal on war debt was an obvious necessity for keeping interest rates in check. But raising taxes was often difficult for administrative reasons (new taxes could not be levied and collected quickly), because higher taxes were ideologically objectionable to the war party, and because higher taxes would, at some point, undermine support for the war. Given those constraints, turning to the printing press was an obvious choice even though it also came with a cost: inflation that also threatened to undermine support for the war provided the public worked out the connection between the inflation and the financial policy of the government. This is a simple story, and often needs elaboration and modification to match particular experiences, but it does a good job of explaining the major features of the American experience from the Revolution to the present day.

Appendices

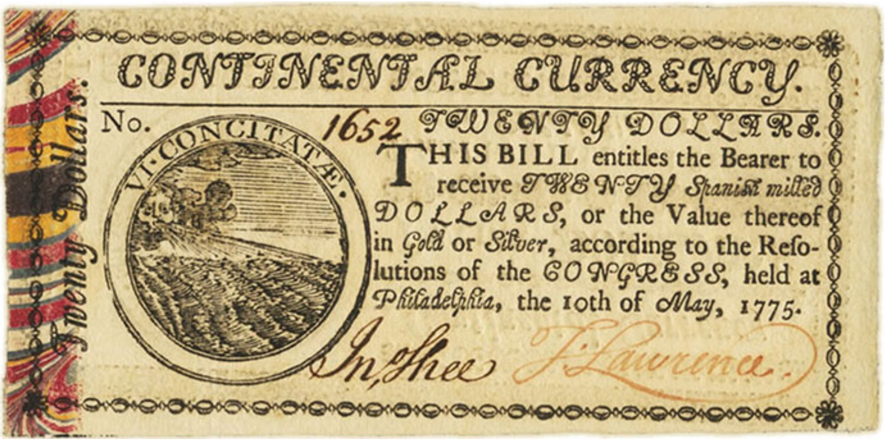
Appendix 1: The Means of Finance for America's Wars

	Revolution	War of 1812	Mexican War	Civil War (North)
Start	April 1775	June 1812	May 1846	April 1861
End	September 1783	December 1814	February 1848	April 1865
Means of finance	1. Printing money	1. Borrowing from the public	1. Borrowing from the public	1. Borrowing from the Public
	2. Borrowing from the public	2. Printing money	2. Land grants	2. Taxes
	3. Borrowing from foreign countries	3. Taxes		3. Printing money
	4. Land grants	4. Land grants		4. Conscription
	5. Conscription	5. Letters of marque		
	6. Letters of marque			
	Spanish-American and Philippine-American War	World War I	World War II	Korean War
Start	April 1898	April 1917	December 1941	June 1950
End	July 1902	November 1918	September 1945	June 1953
Means of finance	1. Taxes	1. Borrowing from the public	1. Taxes	1. Taxes
	2. Borrowing from the public	2. Printing money	2. Borrowing from the public	2. Conscription
	3. Printing money	3. Taxes	3. Printing money	
		4. Conscription	4. Conscription	
	Vietnam War	Persian Gulf War		
Start	August 1964	January 1990		
End	April 1973	March 1990		
Means of finance	1. Borrowing from the public	1. Contributions of foreign governments		
	2. Taxes			
	3. Printing Money			

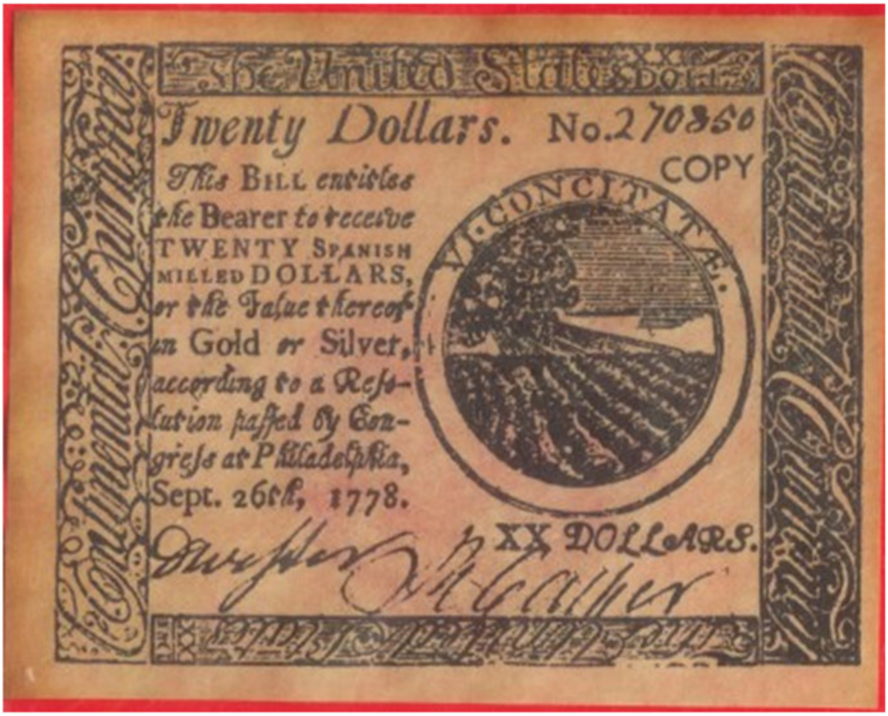
Sources: Dewey (1930, passim), Rockoff (2012, passim), and other financial histories. I have tried to order the means of finance from most to least important, but in some cases, particularly when it comes to nonmonetary forms of finance such as land grants, the ordering is just an informed (hopefully) judgment

Appendix 2: A Catalog of Some of the Currency That Helped Finance America's Wars

1. A Continental note for 20 dollars issued according to the May 1775 resolution of Congress. The resolution promised redemption in 3 Spanish milled Dollars (pesos) between 1779 and 1782. Redemption would be carried out by the states, which would also make the notes legal tenders.

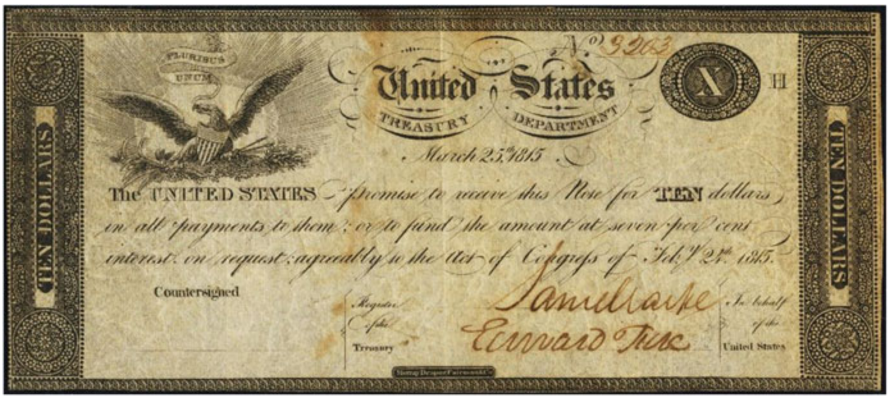


2. A Continental note for 20 dollars issued according to a resolution of September 1778. The note promised redemption (according to Farley Grubb's 2011b, 15, calculations) between 1815 and 1817.



3. A note issued by Virginia in 1776. The presence of state-issued notes and counterfeits – despite the death threat – make it difficult to compute the stock of money during the Revolution.

4. Ten-dollar interest-bearing Treasury Note from the War of 1812



5. A Civil War greenback.

6. A National Bank note issued by the First National Bank of Newark. It was secured by bonds of the United States.

7. A Federal Reserve Note from 1914.

References

Bank, Steven A., Kirk J. Stark, and Joseph J. Thorndike. 2008. *War and taxes*. Washington, DC: Urban Institute Press.

Barro, Robert J. 1987. Government spending, interest rates, prices and budget deficits in the United Kingdom, 1701–1918. *Journal of Monetary Economics* 20: 195–220.

Barro, Robert J. 1989. The neoclassical approach to fiscal policy. In *Modern business cycle theory*, ed. Robert J. Barro. Cambridge, MA: Harvard University Press.

Bockstruck, Lloyd DeWitt. 1996. *Revolutionary war bounty land grants: Awarded by state governments*. Baltimore: Genealogical Pub. Co.

Cagan, Phillip. 1956. The monetary dynamics of hyperinflation. In *Studies in the quantity theory of money*, ed. Milton Friedman. Chicago: University of Chicago Press.

Calomiris, Charles W. 1988. Institutional failure, monetary scarcity, and the depreciation of the continental. *The Journal of Economic History* 48(1): 47–68. March.

Dewey, Davis Rich. 1931. *Financial history of the United States*. New York: Longmans, Green and Co.

Flexner, James Thomas. 1984. *Washington: The indispensable man*. Boston: Little Brown.

Friedman, Milton. 1952. Price, income, and monetary changes in three wartime periods. *The American Economic Review* 42(2), Papers and Proceedings of the Sixty-fourth Annual Meeting of the American Economic Association) (May): 612–625.

Friedman, Milton, and Anna J. Schwartz. 1963. *A monetary history of the United States*. Princeton: Princeton University Press.

Friedman, Milton, and Anna J. Schwartz. 1970. *Monetary statistics of the United States*. New York: Columbia University Press.

Friedman, Milton, and Anna J. Schwartz (eds.). 1982. *Monetary trends in the united states and the united kingdom: Their relation to income, prices, and interest rates, 1867–1975. National bureau of economic research monograph*. Chicago: University of Chicago Press.

Gates, Paul Wallace. 1968. *History of public land law development*. Washington: U.S. Government Publishing Office.

Grubb, Farley. 2008. The continental dollar: How much was really issued? *The Journal of Economic History* 68(1): 283–291.

Grubb, Farley. 2011a. The distribution of congressional spending during the American Revolution, 1775–1780: The problem of geographic balance. In *The spending of States. Military expenditure during the long eighteenth century: Patterns, organization, and consequences, 1650–1815*, ed. Stephen Conway and Rafael Torres Sãnchez, 257–284. Saarbrücken: VDM Verlag Dr Müller.

Grubb, Farley. 2011b. *The continental dollar: Initial design, ideal performance, and the credibility of congressional commitment*. NBER Working Paper No. 17276, Issued in August.

Hibbard, Benjamin Horace. 1965 [1924]. *A history of the public land policies*. Madison and Milwaukee: The University of Wisconsin Press.

Hickey, Donald R. 1989. *The war of 1812: A forgotten conflict*. Urbana: University of Illinois Press.

Historical Statistics of the United States Millennial Edition Online. 2006. Retrieved from http://hsus.cambridge.org/.

Kang, Sung Won, and Hugh Rockoff. 2006. *Capitalizing patriotism: The liberty loans of World War I*. NBER Working Paper 11919.

Keynes, John Maynard. 1920 [1919]. *The economic consequences of the peace*. New York: Harcourt, Brace and Howe.

McCusker, John J. 2001. *How much is that in real money? A historical commodity price index for use as a deflator of money values in the economy of the United States*, 2nd ed. Worcester: American Antiquarian Society.

Meltzer, Allan H. 2003, 2009. *A history of the federal reserve*. Two volumes. Chicago: University of Chicago Press.

Mill, John Stuart, and W. J. Ashley. 1936 [1848]. *Principles of political economy, with some of their applications to social philosophy*. London/New York: Longmans, Green.

Ratchford, Benjamin Ulysses. 1941. *American state debts*. Duke University Publications. Durham: Duke University Press.

Rockoff, Hugh. 2012. *America's economic way of war: War and the US economy from the Spanish-American War to the Persian Gulf War*. New York: Cambridge University Press.

Smith, Adam. 1976 [1776]. *An inquiry into the nature and causes of the wealth of nations*. The Glasgow edition of the works and correspondence of Adam Smith., ed. R.

Truman, Harry S. 1955, 1956. *Memoirs*. 1st ed. Garden City: Doubleday.

Lessons Learned? British Mobilization for the Two World Wars

Stephen Broadberry and Peter Howlett

Abstract We compare the mobilization of the British economy during the two world wars, asking to what extent performance improved in World War II as a result of lessons learned from the experience of World War I. We find that government controls were introduced more quickly and comprehensively in World War II, which improved the scale and speed of mobilization. Better fiscal and financial management also led to less inflation. However, the external account proved more of a problem during World War II, and the greater reliance on planning rather than market mechanisms may have had adverse effects on productivity performance in the postwar world. The setback to national wealth was greater during World War II, but this did not lead to heavier reparation demands, as a result of lessons learned from the Versailles settlement.

Keywords World wars • Mobilization • Lessons • Finance • Planning • Markets • Wealth

Introduction

Although a vast amount has been written about Britain during the two world wars, the economic history of these important episodes remains neglected beyond the official histories commissioned shortly after each war. The official studies after World War II are organized around administrative issues and provide a wealth of detail but little in the way of an organizing framework for understanding the

Draft paper for conference to honor Mark Harrison: Economic History of Coercion and State Formation, University of Warwick, 31 March–1 April 2014.

S. Broadberry (✉)
Nuffield College, Oxford University, Oxford OX1 1NF, UK
e-mail: stephen.broadberry@nuffield.ox.ac.uk

P. Howlett
London School of Economics, London, UK
e-mail: W.P.Howlett@lse.ac.uk

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_8

key economic issues. The histories of World War I were commissioned by the Carnegie Endowment for International Peace and formed part of an international project, but are equally heavy on administrative detail and light on analytical clarity. We attempted to address this gap with two papers on the British economy during World War II (Broadberry and Howlett 1998) and during World War I (Broadberry and Howlett 2005). These papers were part of an international comparative project initiated by Mark Harrison to compare the major combatant economies in the two world wars. In writing these papers, we were therefore conscious of the need to focus on contrasts between Britain and the other major combatants in each war considered separately. However, by adopting a similar framework of analysis for each war, we laid the foundations for this chapter, which provides a study comparing the performance of the British economy during the two world wars.

We compare the two war economies, asking to what extent performance improved in World War II as a result of lessons learned from the experience of World War I. The main areas covered are: (1) the scale of mobilization; (2) fiscal and financial management; (3) managing the external account; (4) the reliance on planning versus the market mechanism; and (5) the impact on wealth.

The Scale of Mobilization

War Spending

We begin our analysis of the British economy by examining the path of real GDP in World War I compared to that in World War II. Feinstein (1972: Table 6) provides separate estimates based on the expenditure and income sides of the national accounts, which he averages to produce a compromise estimate of real GDP, and it is this which is reported in Table 1. The general pattern of real GDP growth was similar in both wars, expanding until reaching a peak after 5 years and then declining. However, whereas in World War I the compromise estimate of real GDP

Table 1 Real GDP of the UK at constant factor cost, (% of prewar year)

World War I (1913 = 100)		World War II (1938 = 100)	
1913	100.0	1938	100.0
1914	101.0	1939	101.1
1915	109.1	1940	111.1
1916	111.5	1941	121.2
1917	112.5	1942	124.2
1918	113.2	1943	127.0
1919	100.9	1944	121.9
		1945	116.6
		1946	111.5

Source: Feinstein (1972: Table 6)

Table 2 Government expenditure as a share of GDP at constant market prices (%)

World War I (1913 = 100)		World War II (1938 = 100)	
1913	8.1	1938	13.5
1914	11.5	1939	19.6
1915	31.2	1940	39.9
1916	35.6	1941	47.2
1917	38.7	1942	49.0
1918	37.7	1943	49.7
1919	18.1	1944	48.8
		1945	42.2
		1946	23.3

Source: Feinstein (1972: Table 5)

rose to a peak in the last year of the war, which was 13.2 % above the 1913 level and then dropped back close to the 1913 level when the war ended, in World War II the peak in 1943 was 27 % above the prewar level (that is, rising by more than twice the rate achieved in World War I) and in 1946 was still 11.5 % above the prewar level. It should also be noted that in 1938 real GDP was 27.6 % higher than in 1913. The growth of real GDP in wartime was a significant factor in the financing of the war effort; for example, Harrison (1988: 185) estimated that the growth of real GDP in World War II helped to provide more than half of the domestic finance for war expenditure.

War is waged by the state and therefore one simple measure of wartime mobilization is the increase in state expenditure as a percentage of GDP. This is shown in Table 2, again drawing on Feinstein (1972: Table 5). In both conflicts there was a very rapid expansion in government expenditure in the first 2 years of war: it almost quadrupled as a percentage of real GDP in World War I and almost trebled in World War II. However, there was a 10-percentage-point difference in the peaks of the share of government expenditure, it being higher in World War II where for 4 years (1941–1944) almost half of GDP was accounted for by state expenditure. The government share did fall dramatically in the first year of peace (and more sharply in the case of World War II), but perhaps coincidentally in both cases the share in the immediate postwar year was approximately 10 percentage points higher than the share in the immediate prewar year. In both wars the expansion in government expenditure came primarily at the expense of consumption expenditure, but the bite of wartime austerity was much deeper in World War II: the share of consumption expenditure fell from 77.2 % in 1913 to a low of 60.2 % in 1917 compared to a decline from 78.8 % in 1938 to a low of 51.9 % in 1943; put another way, the consumption share in 4 years during World War II was less than the low point in World War I (Broadberry and Howlett 1998: 47, 2005: 210). Despite this, when taking account of real economic growth and population the situation looks different: when measured in constant £1913 consumption expenditure per head fell from £45 in 1913 to £38 in 1917 whereas in 1938 it was £54 and fell to a low of £43 in 1943 (Feinstein 1972: Table 17).

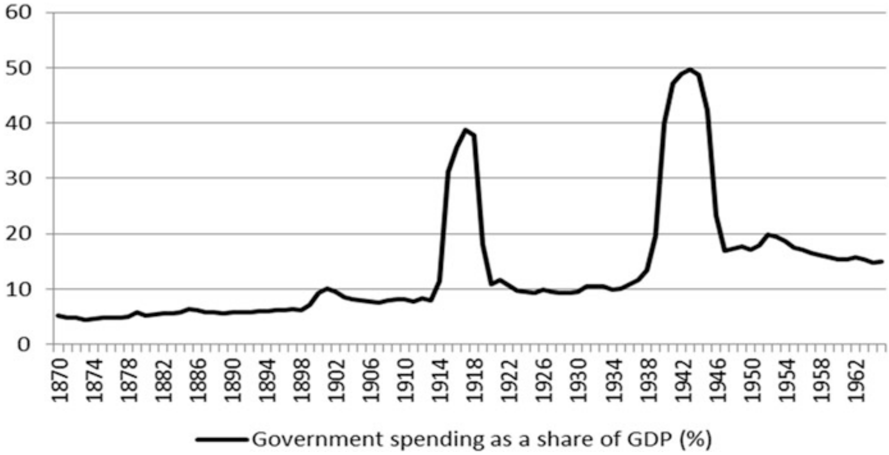


Fig. 1 UK government spending as a share of GDP at constant prices (%) (Source: Feinstein (1972: Table 5))

Figure 1 demonstrates the unprecedented scale of the surge in government spending during World War I, which was dramatically higher than that seen during the Boer War at the turn of the century, and only slightly lower than during World War II. It is easy to understand why World War I has been seen as the first "total war" (French 1982).

Labor Mobilization

Another measure of mobilization for war is the proportion of those employed who are drafted into the armed forces, as shown in Table 3. The expansion of the armed forces followed a similar trajectory in both wars—the more rapid expansion in 1914 compared to that in 1939 can be explained by the fact that World War I began in July 1914 whereas World War II began in September 1939. However, if we consider the level of total employment, its 1918 peak was 5.8 % larger than it had been in 1913, whereas in World War II the 1943 peak was 16.6 % above its 1938 level. Thus there were almost 3.5 million more people in the total employment peak in World War II compared to World War I and approximately 700,000 more in the Armed Forces (Feinstein 1972: Table 57). One reason for the larger expansion in total employment compared to the prewar level in World War II was that the pool of unemployment was much larger in 1938 compared to 1913. From a mobilization perspective the size of the working population is a better indicator than total employment, and here there was not a significant difference: the World War I peak in 1918 was 5.2 % larger than it had been in 1913, whereas in World War II the 1943 peak was 6.4 % above its 1938 level (Feinstein 1972: Table 57). In both wars, the increase in the size of the

Table 3 Civilian employment and armed forces as % of total employment

	World War I		World War II		
	Civilian employment	Armed forces		Civilian employment	Armed forces
1913	98.0	2.0	1938	98.0	2.0
1914	96.0	4.0	1939	97.8	2.2
1915	88.1	11.9	1940	90.2	9.8
1916	83.5	16.5	1941	85.9	14.1
1917	80.1	19.9	1942	83.5	16.5
1918	79.4	20.6	1943	80.9	19.1
1919	89.9	10.1	1944	79.8	20.2
			1945	78.8	21.2
			1946	88.1	11.9

Source: Feinstein (1972: Table 57)

working population was partly due to population increase, but the main factor was the mobilization of women (Ministry of Munitions [1923], Vol. VI, part IV; Howlett 2004: 18–19).

Lessons Learned

The scale of mobilization was very high during World War I, certainly when compared with previous experience. However, it was substantially higher again during World War II: for example, the peak share of government spending in GDP during World War II was 49.7 % in 1943, more than 10 percentage points higher than the World War I peak of 38.7 % in 1917.

There were significant differences between how the state approached the war economy in both conflicts. For example, in World War I the government wished to maintain "business as usual", a principle that was gradually chipped away as the war became more protracted and extensive, whereas the British state explicitly abandoned this principle with regards to the rearmament program as early as March 1938 (Hancock and Gowing 1949: 70; see also the section on Planning versus the Market below). Indeed, the ability to learn from the experience of World War I was crucial to the transition of the economy in World War II.

During World War I, Britain had to face for the first time the economic and social dilemmas posed by fighting a total war. It was an administrative learning experience, at times a painful one, whereby state controls were introduced in a piecemeal fashion (Tawney 1943). Crucially, however, the British state used the experience of World War I to draft plans in the 1930s which could be implemented in the event of another large-scale conflict. Although such extensive pre-war preparations did not mean that state management and control of the economy in the war proceeded without problems, it greatly reduced the administrative friction and disruption caused by moving from a peacetime economy to a war economy, which in turn helped with gaining public acceptance of the measures taken. For example, by the time Britain declared war on Germany in September 1939, 50 million ration books were ready for issue (Zweiniger-Bargielowska 2000: 16–17).

Many of the measures introduced in the first 18 months of World War II had been adopted or were refinements of measures first adopted in World War I, and one of the most important lessons that was learned from that experience, by the state, capital, and labor: the need for the state to evolve a coordinated and comprehensive management of the wartime economy (Hancock and Gowing 1949: 45–72). The Emergency Powers (Defence) Act, passed on 24 August 1939, was an enabling act that gave the British state extensive immediate powers and the potential to extend them if necessary. Significantly, this bill was passed before Britain declared war on Germany. Initially financial planning held center stage in wartime planning, necessitating a strong export drive to pay for vital war stores being bought in the United States and for increasingly expensive imports (Sayers 1956: 257). However, when Churchill became prime minister in May 1940 financial planning was replaced by physical planning and the Treasury was displaced as the most important state body by a series of war cabinet committees concerned with the allocation of materials, labor, and other physical resources (Scott and Hughes 1955; Hancock and Gowing 1949: 88–95; Howlett 1993: 361–378).

Fiscal and Financial Management

War Finance

The exceptional nature of the expansion in government expenditure in both wars has already been noted, and this in turn required the state to engage in exceptional fund-raising exercises. Generally the state can raise funds by increasing taxation, increasing borrowing, or printing more money; and, as shown by Table 4, during both wars the British state did all three.

The state increased its tax take using very similar methods in both conflicts, such that in both wars there was a marked relative shift away from indirect taxation to direct taxation (Broadberry and Howlett 2005: 215–217; Howlett 2004: 13–15). Income tax revenue was boosted by raising the rate of tax and by pulling more people into the tax net, either directly by lowering the exemption limit or indirectly via inflation.

In World War I the standard income tax rate was doubled to 12 % in November 1914, and was then raised progressively throughout the war, finally reaching 30 % in 1918–1919. The exemption limit was reduced from £160 to £130 in 1915, which combined with wage and price inflation to increase the number of taxpayers from 1.1 million prior to the war to 3.5 million in the final year of the war (Mallet and George 1929: 322–328, 395–398). Most of these new taxpayers were wage earners who became liable for tax between 1916 and 1918 (Balderston 1989: 236–237). In World War II the standard rate of income tax doubled from 25 % in 1937–1938 to 50 % in

Table 4 Financing the UK central government deficit (£m)

	Total revenue	Total expenditure	Budget deficit	Increase in: Domestic long debt	Domestic short debt	Money base	Other finance
1913/1914	198	197	–1				
1914/1915	227	561	334	391	64	73	−194
1915/1916	337	1559	1222	458	510	27	227
1916/1917	573	2198	1625	1477	95	56	−3
1917/1918	707	2696	1989	748	484	42	715
1918/1919	889	2579	1690	1019	247	123	301
1938	673	781	108	77	−179	18	192
1939	771	1261	490	72	280	18	120
1940	1158	3273	2115	983	517	70	495
1941	1905	4727	2822	1650	903	109	160
1942	2314	5223	2909	2100	476	191	142
1943	2759	5585	2826	1955	1.017	200	−346
1944	2897	5569	2672	1711	1081	190	−310
1945	2806	4937	2131	1885	557	184	−495

Sources: World War I: Morgan (1952: 98, 107); World War II: Central Statistical Office (1951: 202); Capie and Webber (1985: Table 1.1)

1941–1942 (Sabine 1970: 304). More importantly, the proportion of the population paying tax was greatly widened through such measures as the introduction in 1943 of the Pay-As-You-Earn scheme (Sayers 1956: 112–113).

A significant wartime fiscal innovation in World War I was the Excess Profits Duty and it was again used in World War II. It was the first tax to be levied on companies as opposed to their shareholders. Introduced in the September 1915 budget it taxed profits in excess of a stipulated peacetime standard. The rate was initially 50 % but was increased to 60 % in April 1916 and then 80 % in May 1917. In World War II the Excess Profits Tax was initially set at a rate of 60 %, and raised in 1940 to 100 % (Sabine 1970: 158–159; 168–169). There is no doubt that it was subject to much evasion and fraud in both wars but even so it was spectacularly successful as a revenue generator: by 1918–1919 it was raising £285 million for the exchequer, almost a third of total revenue, making it the single most important tax wielded by the state; in World War II it was relatively less important, but at its peak in 1943 it still generated £482 million (Broadberry and Howlett 2005: 217; Sayers 1956: 223).

Despite the impressive increase in government revenue, Table 4 shows that in both conflicts the onset of war quickly overwhelmed revenue capacity: in 1914–1915 revenue funded only 40 % of expenditure and in 1940 only 35 %. Hence, the government had to turn to other sources of finance to cover the budget deficit. In both wars long-term domestic borrowing was the most significant factor: in World War I about 60 % of the deficit was financed this way, and in World War II the proportion was slightly higher, financing about two-thirds of the deficit. Short-term floating debt, principally in the form of Treasury Bills and Treasury Deposit Receipts was another significant source of financing the deficit (Kirkaldy 1921: 153–162; Sayers 1956: 223). In World War I another important source of finance was borrowing from abroad, particularly from the United States (Kirkaldy 1921: 175–183). Of course, the United States was even more important to the war effort in World War II via Lend-Lease aid, which is not captured in Table 4.

To a limited extent, the government also financed the deficit by allowing an inflationary expansion of the money base, more so in World War I (Capie and Wood 1994: 232–234). Goodhart (1986) sees the sharp increase in the money base (M0) during the first few months of World War I as necessary to meet a run to cash by UK residents. However, historians generally agree that the injection of liquidity was too large and went on for too long, and was thus a contributing factor to wartime inflation (Capie and Wood 1994: 233–234). There has been no formal attempt to measure the success of anti-inflation policy during World War I along the lines of Capie and Wood's (2002) study of World War II. However, Table 5 shows that the GDP deflator, the retail price index, and the money supply (measured by broad money, M3) all approximately doubled between 1914 and 1918. Between 1939 and 1945, by contrast, although the money supply approximately doubled, the GDP deflator and the retail price index increased only by around 50 %.

Table 5 UK money and prices (% of prewar year)

	M3	GDP deflator	Retail price index
1913	100	100	100
1914	108	101	101
1915	125	112	121
1916	138	127	143
1917	156	161	173
1918	190	191	199
1919	232	225	211
1938	100	100	100
1939	99	104	103
1940	109	113	117
1941	126	124	129
1942	142	133	137
1943	162	139	142
1944	184	147	145
1945	209	151	148

Sources: Capie and Webber (1985; Table 1.3); Feinstein (1972: Tables 61, 65)

Lessons Learned

Whereas we have noted that there were some broad similarities in terms of how the state raised revenue during the two world wars, there was a significant difference in the ethos of fiscal policy in the two conflicts. In World War II the state took an earlier and more explicit approach to managing the financial resources of the economy, mainly to better control inflation.

At least until 1917, British fiscal policy in World War I was governed by the "McKenna Rule", which saw the duty of fiscal policy as raising enough revenue to pay for normal peacetime expenditure plus the interest on war loans (French 1982: 106). This policy has been criticized for being too cautious and for stoking wartime inflation (by not mopping up excess expenditure in the economy). However, it has also been argued that political, social, and practical constraints meant that it would have been difficult for the state to pursue a more vigorous policy (Peden 1985: 40–44; Balderston 1989: 222–224).

At the heart of the new approach in World War II was an economist, John Maynard Keynes. The traditional account usually places Keynes's contribution to the conduct of fiscal and monetary policy close to the center of the story (Sayers 1956; Pollard 1992). Keynes developed the idea of an "inflationary gap" to analyze the problem of war finance (Keynes 1939). He viewed the orthodox "Treasury View" of calculating how much tax revenue would be available on the principle of how much people would be willing to pay as a recipe for inflation. He argued, rather, that the government needed first to calculate national income, so as to assess the war potential of the economy, and then set taxes at the level needed to bring about the necessary transfers from the taxpayers to the government. The extra wartime taxes could be treated as forced savings or deferred pay to be repaid after the war. This had the additional advantage of building up potential purchasing power that could be released in the event of a postwar slump, as well as financing the war effort. To the extent that the government failed to achieve the required levels of taxation or forced savings, there would be an inflationary gap, because the excess of aggregate demand over aggregate supply would bid up prices. In making this analysis, Keynes was influenced by his work in the Treasury in World War I and in *How to Pay for the War* he made explicit reference to the inflationary experience of World War I (Keynes 1939: 422–425). The 1941 budget made explicit use of the national accounts and the idea of the inflationary gap: the Keynesian Revolution (albeit in the peculiar setting of total war) had arrived (Broadberry and Howlett 1998: 48–49). While embracing a Keynesian approach to limit demand–pull inflation, this budget also utilized cost of living subsidies to tackle cost–push inflation (Sayers 1956: 90).

Another tool in the state armory for controlling demand was rationing. In World War I rationing was not introduced until 1918, although some localized rationing had begun in November 1917 and eventually covered sugar, meat, butter, margarine, bacon, ham, and lard (Beveridge 1928: 206–207; Barnett 1985: 146). In World War II, however, rationing was used from the start and eventually far more extensively: by the spring of 1945, rationing covered about one half of consumer spending on goods at prewar values and about one-third of consumer spending on goods and services (Mills and Rockoff 1987: 209).

In conclusion, the state was more successful during World War II in controlling the price level, which Capie and Wood (2002) attribute to taxation policy, bond finance and, in contrast to World War I, the widespread use of ration coupons. The inflationary consequences of the expansion of the money base were also muted by the extensive controls exercised over the banking sector, thus limiting the money multiplier effects.

Managing the External Account

The Impact of War on the External Account

There is a marked contrast in the experience of the British economy in the two world wars when we consider the balance of payments (Table 6). In World War I, merchandise exports, while not exceeding their 1913 value, remained relatively stable between 1914 and 1918, whereas in World War II they slumped by almost half between 1939 and 1943. It should be noted, however, that these values are in current prices; export prices increased by 153 % between 1913 and 1918 and by 85 % between 1938 and 1945, so that real exports fell substantially during both conflicts (Feinstein 1972: Table 64). At the same time, merchandise imports rose in both wars in current prices, although this was due to a substantial increase in

Table 6 UK balance of payments (£billion)

	Merchandise exports	Merchandise imports	Merchandise balance	Invisible balance	Net transfers	Current balance
A. Current account						
1914	0.526	−0.696	−0.170	0.315	−0.020	0.125
1915	0.484	−0.852	−0.368	0.395	−0.050	−0.023
1916	0.604	−0.949	−0.345	0.520	−0.050	0.125
1917	0.597	−1.064	−0.467	0.575	−0.080	0.028
1918	0.532	−1.316	−0.784	0.580	–	−0.204
1919	0.963	−1.626	−0.663	0.605	–	−0.058
1939	0.440	−0.840	−0.400	0.200	0.0	−0.2
1940	0.411	−1.126	−0.715	0.215	−0.3	−0.8
1941	0.365	−1.132	−0.767	0.033	−0.3	−1.1
1942	0.271	−0.992	−0.721	0.021	−0.4	−1.7
1943	0.234	−1.228	−0.994	−0.806	−0.3	−2.1
1944	0.266	−1.294	−1.028	−1.372	−0.1	−2.5
1945	0.399	−1.053	−0.654	−0.546	−0.4	−1.6

(continued)

Table 6 (continued)

	Government lending	Government borrowing	Net government lending	Sale of investments	Other transactions
B. Capital account					
1914	–	–	–	–	−0.125
1915	−0.298	0.053	−0.245	0.043	0.225
1916	−0.530	0.319	−0.211	0.110	−0.024
1917	−0.563	0.532	−0.031	0.060	−0.057
1918	−0.297	0.381	0.084	0.023	0.097
1919	−0.137	0.057	−0.080	0.029	0.109
1939	–	–	–	0.0	0.2
1940	–	–	–	0.2	0.6
1941	–	0.3	0.3	0.3	0.5
1942	−0.1	1.2	1.1	0.2	0.4
1943	−0.7	2.1	1.4	0.2	0.5
1944	−0.8	2.6	1.8	0.1	0.6
1945	−0.5	1.3	0.8	0.1	0.7

Sources: World War I: Morgan (1952: 304, 341); World War II: Central Statistical Office (1951: 142); Sayers (1956: 495, 499)

import prices, with imports in constant prices falling. The overall result was a deficit on the balance of trade measured in current prices in every war year. However in World War I this did not lead to a current account deficit in most war years due to the resilience of invisible earnings, which rose from £315 million in 1914 to £580 million in 1918; the overall current account for the period 1914–1918 was just in surplus. In contrast, in World War II invisibles were much weaker and from 1943 exacerbated the merchandise deficit, with the result that the current account was in deficit in every year and the overall deficit for the period 1939–1945 was £10 billion.

The difference between the two conflicts can also been seen in the capital account. For example, the fact that in World War I the external account was not a serious threat to the war effort meant the government felt confident enough to loan more to allies than it borrowed from them in all years apart from 1918. Total overseas borrowing by the government during the war amounted to £1365 million by the end of the financial year 1918–1919, with 75 % coming from the United States, but this was more than offset by wartime government overseas loans which, by the end of the financial year 1918/19, totaled £1741 million (Morgan 1952: 317, 320–321). In contrast, in World War II government borrowing exceeded government lending in each year and net borrowing, primarily from the United States, for the war period amounted to £5.4 billion. In World War II the external situation deteriorated far more rapidly and with potentially far more serious consequences than in World War I: external liabilities more than doubled between December 1939 and December 1941, by mid-1940 assets in North America were being sold off cheaply in a desperate attempt to pay for American goods, and by the beginning of 1941 hard currency reserves had been exhausted (Sayers 1956: 438–464). To a large extent the situation was rescued by the passing of the Lend-Lease Act in the United States in March 1941; this would prove to be the single most important method of financing the current account external deficit in World War II. The introduction of lend-lease considerably relaxed the external constraint and allowed a much greater degree of specialization by Britain on war work than would otherwise have been possible (Allen 1946). The deficit was also covered by the £1.1 billion sale of investments (the equivalent figure for World War I was £236 million) and the accumulation of £3.4 billion of liabilities (Broadberry and Howlett 1998: 52–53).

Although Britain was effectively off the gold standard during World War I, the authorities did attempt to keep sterling at the prewar parity of $4.86. However, the pound depreciated during 1915, reflecting the deterioration in the trade balance, reaching a low of $4.49 in October. The entry of the United States into the war saw the exchange rate recover to $4.76, where it more or less remained until Britain formally left the gold standard in April 1919 (Pollard 1992: 27). In World War II, despite the massive current account imbalance, the exchange rate was maintained at a fixed parity of $4.03, about 20 % below the old gold standard parity, protected by a system of import controls and foreign exchange restrictions (Pollard 1992; 178).

Lessons Learned

From the perspective of the external account, the issue is not primarily "lessons learned" but one of legacy: Britain's experience in World War I weakened its long-term international position and that in turn meant that behavior was more constrained during World War II.

In 1914 central gold reserves were £34 million, other monetary gold stood at £123 million and dollar securities totaled £535 million (Pollard 1992: 27). However, World War I was a watershed for the international economy, and the central role of Britain in the pre-1914 world economy was lost (Wrigley 2000). The problems for the British economy were to be long term: the sale of overseas assets, the postwar external changes which exposed the wartime overseas borrowing policy, and, it is argued, the inability to defend the value of sterling weakened the external position of the economy in the interwar period and saw supremacy in international trade and finance pass to the United States (Burk 1985). Thus, while it was the rise in domestic debt which dominated the dramatic rise in the national debt during the war (less than a fifth of the national debt of £7280 million in March 1919 was accounted for by foreign debt) the weakening of Britain's international situation, which was a direct consequence of the war, did reduce the capacity of the economy to service the debt in the interwar period.

Hence, whereas the balance of payments position permitted the British government to act as a net lender to the Allies during World War I, a substantial current account deficit during World War II made the British government a major net borrower on capital account. Perversely, however, loan defaults after World War

I put significant pressure on the interwar British economy, whereas the massive British borrowing during World War II had a less severe economic impact in the medium term because of the forgiving of American Lend-Lease aid.

Planning Versus the Market

The Growing Role of Government Controls During World War I

We have previously noted that the government was slow during World War I to appreciate the need for large-scale intervention and coordination when fighting a total war. However, it would be wrong to characterize the economy in the early years of the war as operating as if peacetime conditions still held. It was not "business as usual" because from quite early on the state was intervening in markets and the war was encroaching on normal economic practice. However, state intervention in and management of the economy was relatively ad hoc in approach until 1917 and tended to be reactive rather than proactive (Lloyd 1924; Broadberry and Howlett 2005: 222–224). The spread of government controls was generally slow, because the economic and material burden of the war was initially underestimated. Prewar plans had envisaged a strategy based on naval blockade with an army of about 130,000 troops, plus the financing of European allies (Ministry of Munitions 1923, Vol. I, part I: 7–45). The rapid expansion of the armed forces therefore initially overwhelmed the capacity of the economy to equip them, although Trebilcock (1975) doubts whether even an army of 130,000 could have been equipped. Until Lloyd George became Prime Minister in December 1916, intervention in the economy was for very specific purposes; there was no attempt before that date for the state to take general control of the economy.

The most significant embodiment of the spread of government influence was the creation on 9 June 1915 of the Ministry of Munitions with a key role in the co-ordination of war production (Ministry of Munitions 1923; Wrigley 1982). This had two main functions: to supply munitions and stores to the Army and the Admiralty, and to control the supply of materials that were deemed crucial to war production. The Ministry was given wide powers and was not constrained by financial controls from the Treasury. The government softened the blow to the private sector by recruiting many prominent businessmen to run and advise the Ministry. Indeed, businessmen were co-opted by the state in many other areas, so that although the state was displacing the market, it was not necessarily displacing business. In this sense, there was still "business as usual".

Even though government intervention in the economy was extensive by the end of the war, it spread at a slow pace until 1917. Although there were internal and external controls on capital, the control of labor was quite limited compared to the experience of World War II. Indeed, even army conscription was not introduced until March 1916. The government did try to placate labor by negotiating a deal on industrial arbitration and dilution in 1915 and by appointing the trade union leader

John Hodges as a Minister of Labour in 1916. The state built its own factories, the National Shell factories, and took control of the railways, shipping, collieries (from December 1916), flour mills (April 1917) and the Irish distilleries (May 1918) as well as 125 other privately owned factories. It requisitioned the output of several industries (such as jute, flax and glycerine) or used its powers to restrict output or distribution in many other industries (including building, cotton spinning, beer, sugar, timber, fertilizer, iron and steel, and paper) via licensing or by regulating the amount of materials or labor allocated to the industry. It became the main, or only, purchaser of important raw materials (such as sugar, meat, imported wheat, wool, jute, indigo, Russian flax and Italian sulfur) while price fixing was used to restrict war profiteering (Morgan 1952: 46–57; Lloyd 1924).

As with most government intervention, policy in the area of food was reactionary. By the end of 1916 growing shortages and rising prices were causing domestic unrest. This led to the gradual expansion of state control over domestic food production and imports such that by the end of the war the Ministry of Food was responsible for 85 % of the food supply (Beveridge 1928: 57). Although some localized rationing had begun in November 1917, rationing was not introduced at a national level until 1918, but eventually covered sugar, meat, butter, margarine, bacon, ham and lard (Beveridge 1928: 206–207; Barnett 1985: 146). Differential dietary requirements were met by bread, which had been subsidized since September 1917 and was freely available (Zweiniger-Bargielowska 2000: 12–13).

Controls During World War II

During the 1930s governments made extensive preparations to move more quickly to a total war footing, in the belief that controls had been adopted too slowly and on an ad hoc basis during World War I (Hancock and Gowing 1949: 45–72). During the war, in addition to the macroeconomic measures to close the inflationary gap discussed in the section on Fiscal and Financial Management, the government used a barrage of microeconomic measures to ensure that the demand for individual goods was brought into line with supply, including: (1) overall central planning to set priorities; (2) rationing to curtail consumer demand; (3) production quotas and the concentration of production in large units in civilian industries (4) central manpower budgeting to allocate labor across sectors; and (5) central allocation of scarce resources such as steel and capital (Wiles 1952: 125–158).

Although there were mechanisms of control and planning during World War I and during rearmament, Wiles (1952) argues that rational overall planning only really began with World War II. The War Cabinet discussed strategic issues, and overall priorities were fed through a production committee to the supply departments, although the details of the structure changed during the course of the war (Howlett 1993: Chester 1951). At the departmental level, new Ministries of Supply, Home Security, Shipping, Food, Economic Warfare, and Information quickly appeared at the outbreak of war, reflecting the changed priorities of the war economy (Hopkins

1952: 1–4). In the formulation and co-ordination of economic policy, the newly established Economic Section of the War Cabinet Secretariat and the Central Statistical Office came to play an increasingly important role (Chester 1951: 14–19; Cairncross and Watts 1989).

A number of items were rationed from the outbreak of war and rationing gradually spread to more consumer goods and services (Zweiniger-Bargielowska 2000: 9–59). By the spring of 1945, rationing covered about one-half of consumer spending on goods at prewar values and about one-third of consumer spending on goods and services (Mills and Rockoff 1987: 209). Initially, rationing operated on a coupon basis, with consumers entitled to fixed amounts of rationed items (Hancock and Gowing 1949: 446). From 1941, however, a more flexible points system was introduced, whereby coupon points could be spent on a limited number of goods, thus allowing consumers some scope for substitution in line with preferences (Hancock and Gowing 1949: 329–332; Reddaway 1951: 182). It has been argued that the rationing system operated more effectively in Britain than in other countries. Although some writers see this as reflecting a greater spirit of voluntary compliance in Britain, Mills and Rockoff (1987) attribute it mainly to the greater scale of resources devoted to the issue, with a fuller array of controls backed up by both financial and legal resources, ensuring a strict supervision of both production and distribution.

Much civilian production was cut back severely at the beginning of the war, particularly through Limitation of Supplies Orders (Wiles 1952: 151; Hancock and Gowing 1949: 117–118). In many consumer industries, the state also implemented a temporary wartime concentration of production drive to gain economies of scale and standardization. Figures from the Federation of British Industries suggest that this drive released 255,900 workers and 61.2 million square feet of capacity for munitions and related industries (Howlett 1994: 144).

Wartime labor mobilization was helped by the elimination of the mass unemployment of the 1930s. However, the overall civilian labor supply needed to be increased in order to replace males recruited into the armed forces, and this was achieved primarily by raising female participation. Furthermore, the civilian labor supply had to be reoriented away from industries producing inessential civilian items to industries producing war supplies, while maintaining employment and output in essential non-war industries such as fuel and power. Although during the early stages of the war labor problems appeared mainly in the form of bottlenecks with skilled labor, as time went on the general supply of labor was seen as a constraint. From December 1942, with the first Manpower Budget, the problem of the allocation of labor between the production programs of the different government departments was tackled directly (Hancock and Gowing 1949: 146). “Manpower” was the term coined in that bygone age, less gender-conscious than our own, but in wartime the most rapidly growing element was “womanpower”. The government had wide powers of labor compulsion which it used to control the supply of labor to both the armed forces and industry, although where possible it relied on voluntarism and co-operation (Robinson 1951: 50).

Other inputs of vital materials and capital were also controlled by the government. For vital materials such as steel, each government order or licensed private order for a product requiring steel carried with it a right to the required amount of steel. This right, known as an "M form" could be cashed at a steelworks. This was administratively complex and led on occasions to "coupon inflation" when too many M forms were chasing too little steel (Wiles 1952; 148–149). The Capital Issues Committee controlled all new issues on the capital market, but this is not the same as control over physical investment. Although building, and at times machine tools, were subject to close control, most investment was controlled only indirectly through the controls on labor and materials (Wiles 1952: 144).

Markets, Planning, and Economic Performance

The standard approach to the economic history of Britain during the two world wars, reflected in the literature surveyed above, has been to stress the limitations of reliance on market forces, the slowness of governments in World War I to learn that lesson, and the benefits of the swift transition to a planned economy during World War II. Is it possible, however that the lessons were learned too well in Britain, and that the belief in the efficacy of government controls went too far? And could this be a factor in the relatively poor performance of the British economy during the postwar period? To answer the first question, it is necessary to consider the role of market forces in the successful conduct of the British war economy. And to answer the second question, we need to consider the impact on Britain's productivity performance of the restrictions on competition which were consolidated during World War II, and continued into the postwar period, 1945–1979.

Few historians are likely to be persuaded that the achievements of the British war economy can be put down to the smooth operation of market forces during the war itself. But did policymakers underestimate the positive effects of Britain's liberal politico-economic inheritance compared to those of her main rivals, and therefore overestimate the contribution of government intervention and planning? Britain was, along with the United States, the most developed market economy in the world in the first half of the twentieth century and had a political, administrative, and financial history that strengthened her ability to wage war successfully. Olson (1963: 73–116) has made this point strongly in discussing food supply. Prior to World War I, Britain was far more dependent than Germany on imported food supplies and during the war, Germany waged a (militarily) successful submarine campaign to disrupt and destroy British food imports. But the campaign did not succeed in starving Britain to surrender. Olson argues that this was because Britain's prewar free trade policy had greatly reduced the size of the agricultural sector, which in turn gave it a capacity for substitution and flexibility that allowed farmers to respond to the German blockade. Also, unlike Germany, which had boosted its agricultural sector to provide a defense against potential wartime blockade, Britain had not attempted to allow strategic motives to distort its economic advantages in those years. Finally, when the food situation did deteriorate in the war and state intervention became necessary, Olson argues that "its relatively unified electorate and generally efficient civil service" allowed Britain to impose controls and execute them effectively.

In a similar vein, Balderston (1989: 224) argues that the development of London as the leading financial center in the world, and the capacity of the capital market to absorb public debt was extremely important for the British war economy. It provided an efficient mechanism for financing the war effort and acted more generally as "a powerful stabilising agent on the short-term behaviour of the British economy". The inheritance of a strong market economy, together with the financial clout of the City interacted together with a strong public administration and (for the time) a well-developed democratic accountability, to provide an economic and political capacity and flexibility that would help to ensure victory. A comparison between Britain and Germany brings Britain's advantage into sharp relief (Olson 1963; Ferguson 2000). It is important not to be mesmerized by Germany's rapid industrialization from the mid-nineteenth century on the basis of protectionism, state intervention, and universal banks (Gerschenkron 1962). Britain's steadier, more market-oriented development made for a more flexible economy which was better able to stand the strains of total war.

The generally positive evaluation of economic planning during World War II reinforced a disenchantment in some quarters with reliance on market forces that had grown out of the mass unemployment of the Great Depression. Although the Labour Party, which formed a majority government for the first time in 1945, rejected a wholesale move to a planned economy in favor of a mixed economy with an emphasis on the achievement of full employment through Keynesian demand management, there remained in government circles a distrust of competitive market forces, which permeated economic policy (Broadberry 2002). A number of important industries were nationalized, including coal, steel, and the railways, while in other industries restrictions on competition which had been strengthened during the war were allowed to continue as a result of "light-touch" competition policy (Broadberry and Crafts 1996). Broadberry and Crafts (2003) argue that these policies were damaging for Britain's productivity performance during the postwar period, lasting until the change of regime beginning with the first Thatcher government.

The Impact on Wealth

The Accounting Framework

Broadberry and Howlett (1998) developed an accounting framework for evaluating the long-run impact of war on wealth, which they applied to the case of Britain during World War II. They then applied the same framework to Britain during World War I, making possible a comparison of the two wars (Broadberry and Howlett

2005). The first important distinction is between stocks and flows in the system of national accounts. Issues concerned with the scale of mobilization, which have been dealt with in the preceding sections, are best tackled by looking at flows of income, expenditure, and output, and asking what proportion of those flows was devoted to the war effort. However, the long-run impact of the war can best be assessed by looking at the effects on national wealth, defined here to include human as well as physical capital, intangible as well as tangible capital and net overseas assets (Goldsmith et al. 1963; Revell 1967; Kendrick 1976).

Tangible physical capital is the conventional form of capital, consisting of buildings, equipment, and inventories. Intangible physical capital is cumulated expenditure on R&D, which is seen as improving the quality of the tangible physical capital. Tangible human capital is the spending required to produce an uneducated, untrained worker, i.e., basic rearing costs. Intangible human capital is mainly spending on education and training to improve the quality of the human capital, although it also includes other items such as spending on health and safety and mobility costs. In an open economy, the impact of the war on net overseas assets must also be taken into account.

We believe that this accounting framework deals with the main objections of writers such as Hardach (1977: 286) and Milward (1984: 9–27) to previous attempts to quantify the impact of war on the economy. In particular, note that: (1) a clear distinction between stock and flow concepts is maintained throughout; (2) all nominal values are converted to a constant price basis so that values for different years can be added together; (3) human capital calculations take account of the fact that people consume as well as produce; (4) the fact that postwar birth rates rise does not alter the fact that the human capital embodied in those killed by warfare is lost, and this has a negative impact on national wealth as much as any destruction of physical capital, which is usually followed by increased investment to make good war losses; (5) technological change stimulated by the war can be seen as having a positive impact on intangible physical capital, and can be captured by cumulating any increase in R&D above the prewar level; and (6) social spending stimulated by the war can be seen as having a positive impact on intangible human capital, and can be captured by cumulating the increase in social spending above the prewar level.

The Impact of the World Wars on Britain's Wealth

Table 7 presents an assessment of the effects of both World Wars on Britain's wealth using the Broadberry and Howlett (1998) framework. Note that although the values are in constant prices for both wars, they are in 1913 prices for World War I and 1938 prices for World War II. Comparison therefore needs to be made in terms of the percentages of national wealth in the last two lines of Table 7.

Dealing first with losses to physical capital, it is important to include not just losses on land, but also losses arising from the sinking of ships and their cargoes, as well as external disinvestment via the sale of overseas investments, government

Table 7 National balance sheet calculation of the effects of the World Wars on the UK economy

	World War I (£m at 1913 prices)	World War II (£m at 1938 prices)
Physical capital losses:		
On land	360	860
shipping and cargo	384	380
Disinvestment	998	3355
Total physical capital losses	1742	4595
Human capital losses:		
Tangible	88	86
Intangible	58	58
Total physical and human capital losses	1888	4739
Prewar national wealth, excluding human capital	11,682	24,680
Prewar national wealth, including human capital	17,218	38,645
Physical capital losses as % of prewar national wealth, excluding human capital	14.9 %	18.6 %
Physical and human capital losses as % of prewar national wealth, including human capital	11.0 %	12.3 %

Sources: World War I: Broadberry and Howlett (2005: 228); World War II: Broadberry and Howlett (1998: 69–70)

borrowing abroad, and net exports of gold and silver. Expressing these losses as a share of prewar national wealth, defined narrowly in terms of physical capital, yields losses of 14.9 % of wealth during World War I and 18.6 % during World War II. On this basis, World War I was a major setback to national wealth, but not on the same scale as World War II, consistent with the flow data on government spending in Fig. 1.

However, this calculation leaves out losses of human capital, which have also been valued in Table 7, taking account of tangible (i.e., basic rearing costs) and intangible human capital (largely education) embodied in the average British casualty. Since the level of casualties was much higher in World War I than in World War II (755,000 compared with 360,000), the inclusion of human capital in the calculations makes the loss of wealth in the two wars much more equal than if attention is confined to physical capital (11.0 % in World War I compared with 12.3 % in World War II).

In the context of "lessons learned", it is perhaps worth noting that there was much more discussion of the costs of the war after World War I than after World War II. Thus, for example, Bogart's (1920) detailed calculation of the costs of World War I for all the major combatant countries was not repeated after World War II. This can be seen at least partly as a result of a shift in attitudes towards reparations, since the calculation of war costs lends itself easily to the victors presenting a bill to the defeated nations. Since wrangling over reparations was widely seen as a contributory factor to World War II, much less effort was made after 1945 to make the defeated Axis Powers pay.

Conclusions

We conclude briefly by summarizing the similarities and differences between the two world wars and the extent to which lessons learned from World War I were used profitably in World War II:

1. Although the scale of mobilization was very high during World War I, certainly when compared with previous experience, the state built on that experience to mobilize an even greater share of the nation's resources for World War II. In contrast to the slow spread of government controls during World War I, plans were prepared during the 1930s and implemented quickly in 1939.
2. Lessons were also learned in war finance, which was less inflationary during World War II. Although the money supply doubled during both wars, price controls and rationing meant less inflation during World War II.
3. The issue in managing the external account was more one of legacy than lessons learned. Whereas the balance of payments position permitted the British government to act as a net lender to the Allies during World War I, a substantial current account deficit during World War II made the British government a major net borrower on capital account. Perversely, however, loan defaults after World War I put significant pressure on the interwar British economy, whereas the massive British borrowing during World War II had a less severe economic impact in the medium term because of the forgiving of American Lend-Lease aid.
4. The literature on World War I emphasizes the slowness of the government in appreciating the need for large-scale state intervention and co-ordination when fighting a total war. This view is summed up in the memorable phrase "business as usual". A similar tendency to idealize the benefits of state control and to denigrate the achievements of the market appears in the literature on World War II. However, there is a danger in such a view of neglecting the benefits that British planners enjoyed from the inheritance of a liberal market economy. These benefits are most obvious when comparing Britain with Germany during both conflicts. In this case, the lessons of war may have been learned too well, with the state too ready to accept restrictions on the operation of market forces, with adverse effects on Britain's productivity performance
5. The setback to national wealth was greater during World War II than during World War I. However, it makes a significant difference whether or not human capital is included. If attention is limited to physical capital, the scale of the wealth destruction was substantially higher during World War II. However, if human capital is also taken into account, the higher level of casualties during World War I means that the scale of the destruction was more similar. To the extent that calculation of the costs of war led naturally to demands for reparations, one lesson learned from World War I was the need to tread cautiously in this area, given the disastrous consequences of the Versailles settlement.

Archival and Published Primary Sources

Central Statistical Office. 1951. *Statistical digest of the war*. London: HMSO and Longmans.
Ministry of Munitions. [1923] 1976. *History of the ministry of munitions*. Hassocks: Harvester.

References

Allen, R.G.D. 1946. Mutual aid between the US and the British empire, 1941–45. *Journal of the Royal Statistical Society* 109: 243–271.

Balderston, T. 1989. War finance and inflation in Britain and Germany, 1914–1918. *Economic History Review* 42: 222–244.

Barnett, L.N. 1985. *British food policy during the First World War*. Boston: George Allen and Unwin.

Beveridge, W.H. 1928. *British food control*. London: Oxford University Press.

Bogart, E.L. 1920. *Direct and indirect costs of the Great World War*, 2nd ed. New York: Oxford University Press.

Broadberry, S.N. 2002. Economic policy. In *Britain since 1945*, ed. J. Hollowell, 383–403. Oxford: Blackwell.

Broadberry, S.N., and N.F.R. Crafts. 1996. British economic policy and industrial performance in the early postwar period. *Business History* 38: 65–91.

Broadberry, S.N., and N.F.R. Crafts. 2003. UK productivity performance from 1950 to 1979: A restatement of the Broadberry-Crafts view. *Economic History Review* 56: 718–735.

Broadberry, S.N., and P. Howlett. 1998. The United Kingdom: 'Victory at All Costs'. In *The economics of World War II: Six great powers in international comparison*, ed. M. Harrison, 43–80. Cambridge: Cambridge University Press.

Broadberry, S.N., and P. Howlett. 2005. The United Kingdom during World War I: Business as usual? In *The economics of World War I*, ed. S.N. Broadberry and M. Harrison, 206–234. Cambridge: Cambridge University Press.

Burk, K. 1985. *Britain, America and the sinews of power, 1914–1918*. Boston: George Allen and Unwin.

Cairncross, A., and N. Watts. 1989. *The economic section, 1939–1961: A study in economic advising*. London: Routledge.

Capie, F., and A. Webber. 1985. *A monetary history of the United Kingdom, 1870–1982. Volume I: Data, sources, methods*. London: Routledge.

Capie, F., and G. Wood. 1994. Money in the economy. In *The economic history of Britain since 1700, volume 3: 1939–1992*, ed. R. Floud and D. McCloskey, 217–246. Cambridge: Cambridge University Press.

Capie, F., and G. Wood. 2002. Price controls in war and peace: A Marshallian conclusion. *Scottish Journal of Political Economy* 49: 39–60.

Chester, D.N. 1951. The central machinery for economic policy. In *Lessons of the British War Economy*, ed. D.N. Chester, 5–33. Cambridge: Cambridge University Press.

Feinstein, C.H. 1972. *National income, expenditure and output of the United Kingdom, 1855–1965*. Cambridge: Cambridge University Press.

Ferguson, N. 2000. How (not) to pay for the war: Traditional finance and 'Total' War. In *Great War, Total War: Combat and mobilization on the western front, 1914–1918*, ed. R. Chickering and S. Förster, 409–432. Cambridge: Cambridge University Press.

French, D. 1982. *British economic and strategic planning, 1905–1915*. London: Allen and Unwin.

Gerschenkron, A. 1962. *Economic backwardness in historical perspective*. Cambridge, MA: Harvard University Press.

Goldsmith, R.W., R.E. Lipsey, and M. Mendelson. 1963. *Studies in the National Balance Sheet of the United States*, vol. 2. Princeton: Princeton University Press.
Goodhart, A.E. 1986. Comment on 'The Summer of 1914'. In *Financial crises and the world banking system*, ed. F. Capie and G. Wood, 117–119. London: Macmillan.
Hancock, W.K., and M.M. Gowing. 1949. *British war economy*. London: HMSO and Longmans.
Hardach, G. 1977. *The First World War, 1914–1918*. Berkeley: University of California Press.
Harrison, M. 1988. Resource mobilization for World War II: The USA, UK, USSR and Germany, 1938–1945. *Economic History Review* 41: 171–192.
Hopkins, R. 1952. Introductory note. In *Lessons of the British war economy*, ed. D.N. Chester, 1–4. Cambridge: Cambridge University Press.
Howlett, W.P. 1993. New light through old windows: A new perspective on the British economy in the Second World War. *Journal of Contemporary History* 28: 361–378.
Howlett, P. 1994. British business and the state during the Second World War. In *World War II and the transformation of business systems*, ed. J. Sakudo and T. Shiba, 133–153. Tokyo: University of Tokyo Press.
Howlett, W.P. 2004. The wartime economy, 1939–1945. In *The Cambridge economic history of modern Britain: Structural change and growth, 1939–2000*, ed. R. Floud and P. Johnson, 1–26. Cambridge: Cambridge University Press.
Kendrick, J.W. 1976. *The formation and stocks of total capital*. New York: Columbia University Press.
Keynes, J.M. [1939] (1978). "How to Pay for the War". In *The collected writings of John Maynard Keynes, Vol. XXII: Activities 1939–1945, Internal War Finance,* ed. D.E. Moggridge, 41–51. London: Macmillan.
Kirkaldy, A.W. 1921. *British finance during and after the War*. London: Isaac Pitman & Sons.
Lloyd, E.M.H. 1924. *Experiments in state control*. Oxford: Oxford University Press.
Mallet, B., and C.O. George. 1929. *British budgets, second series, 1913/14 to 1920/21*. London: Macmillan.
Mills, G.T., and H. Rockoff. 1987. Compliance with price controls in the United States and the United Kingdom during World War II. *Journal of Economic History* 47: 197–213.
Milward, A.S. 1984. *The economic effects of the two world wars on Britain*, 2nd ed. London: Macmillan.
Morgan, E.V. 1952. *Studies in British financial policy, 1914–25*. London: Macmillan.
Olson Jr., M. 1963. *The economics of the wartime shortage: A history of British food supplies in the Napoleonic War and in World Wars I and II*. Durham: Duke University Press.
Peden, G.C. 1985. *British economic and social policy: Lloyd George to Margaret Thatcher*. Oxford: Philip Allan.
Pollard, S. 1992. *The development of the British economy, fourth edition, 1914–1990*. London: Arnold.
Reddaway, W.B. 1951. Rationing. In *Lessons of the British war economy*, ed. D.N. Chester. Cambridge: Cambridge University Press.
Revell, J. 1967. *The wealth of the nation: The national balance sheet of the United Kingdom, 1957–1961*. Cambridge: Cambridge University Press.
Robinson, E.A.G. 1951. The overall allocation of resources. In *Lessons of the British war economy*, ed. D.N. Chester. Cambridge: Cambridge University Press.
Sabine, B.E.V. 1970. *British budgets in peace and war, 1932–1945*. London: Allen & Unwin.
Sayers, R.S. 1956. *Financial policy, 1939–45*. London: HMSO and Longmans.
Scott, J.D., and R. Hughes. 1955. *The administration of war production*. London: HMSO and Longmans.
Tawney, R.H. 1943. The abolition of economic controls, 1918–1921. *Economic History Review* 13: 1–30.
Trebilcock, C. 1975. War and the failure of industrial mobilisation: 1899 and 1914. In *War and economic development*, ed. J.M. Winter, 139–164. Cambridge: Cambridge University Press.
Wiles, P.J. 1952. Pre-war and war-time controls. In *The British economy*, ed. G.D.N. Worswick and P.H. Ady, 1945–1950. Oxford: Oxford University Press.

Wrigley, C. 1982. The Ministry of Munitions: An innovatory department. In *The war and the state: The transformation of British government, 1914–1919*, ed. K. Burk, 32–56. London: George Allen and Unwin.

Wrigley, C. 2000. The war and the international economy. In *The First World War and the international economy*, ed. C. Wrigley, 1–33. Cheltenham: Edward Elgar.

Zweiniger-Bargielowska, I. 2000. *Austerity in Britain. Rationing, controls, and consumption, 1939–1955*. Oxford: Oxford University Press.

World War II and US Economic Performance

Price V. Fishback and Taylor Jaworski

Abstract One popular view is that World War II ended the Great Depression and led to a postwar boom in the United States. On the other hand, recent studies suggest that the reallocation of resources to meet war demands may have imposed more costs than benefits. In the first part of this chapter we review recent evidence on the war's impact on the United States economy at both the national and local levels. In the second part we extend recent work on the impact of World War II on local economies. We use a simple spatial equilibrium model and data for all United States counties to estimate the long-run impact of the war-related spending on income per capita, population, and median housing values per decade from 1960 to 2010. The empirical results show that the main changes in local economies were due to the reallocation of population toward counties that received more per capita war spending. The War was associated with very little growth in income per capita and median house values. When combined with the model, these results suggest that mobilization for World War II was mostly correlated with an increase in the value of local amenities, which were likely related to the proximity to defense-related industries and associated non-wage benefits.

Keywords United States economic growth • World War II • Defense economics • Migration • Income per capita • Housing values

Introduction

World War II ended the Great Depression in the United States. This view pervades the popular media and national politics (Romer 2011; Krugman 2011). It is also a part of the historical narrative of the "American century" beginning with the post-World War I boom of the 1920s, interrupted by a severe economic downturn in the

P.V. Fishback (✉)
Department of Economics, University of Arizona, Tucson, AZ, USA
e-mail: fishback@email.arizona.edu

T. Jaworski
Department of Economics, Queen's University, Kingston, ON, Canada
e-mail: tjaworski@gmail.com

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*,
Studies in Economic History, DOI 10.1007/978-981-10-1605-9_9

1930s, reignited by mobilization for World War II during the 1940s, and ultimately giving rise to the postwar "golden age" until the 1970s. Traditional measures of economic performance over this period support this narrative. Unemployment decreased and real income per person increased. Increased household saving spurred a postwar boom in consumption; stock prices and corporate profits rose. As the Allied Forces began to win the war, the uncertainty of the Depression years was forgotten and a sense of optimism prevailed. The United States entered the postwar years as a leader economically, politically, and culturally.

On the one hand, this view emphasizes the role of war in sparking changes in politics, economy, and society. The more limited availability of resources imposed constraints that led to more inclusive political and economic institutions or technological change that allowed resources to be used more productively. On the other hand, the reallocation of resources to meet the ends of war may have imposed more costs than the benefits gained from shedding old practices. Moreover, institutional changes have not always led to more inclusion nor has the direction of technological change always led to first-best innovations. This is the lesson of research in the economic history of American warfare in the twentieth century (Goldin 1980; Higgs 1987). This work emphasizes the costs of mobilizing for war, both the fiscal costs and the distortion of incentives that directed resources away from their highest-value uses.

The tension in these two views is between the potential benefits of war in spurring changes in innovation, business, and economic and political institutions, and the costs of the government allocation of resources during wartime and the persistence of these controls after war had ended. The American experience with warfare is unique in that few wars–and none in the twentieth and twenty-first century–were fought within the nation's borders. This has limited the direct loss of domestic, civilian capital and meant the development of the American economy has been less hampered by the destruction faced elsewhere in terms of civilian lives, factories, and cities. Still, the estimated costs associated with mobilization are typically substantial and the impact of World War II on the development of the postwar United States should be understood in terms of the costs and long-run changes to economic and political institutions.

At the local level, the potential benefits of mobilization came in the form of increased productivity as government investment on supply contracts, and new investment exposed workers and firms to new production techniques. In the first part of this chapter we review recent evidence for the impact of World War II on the performance of the United States economy. At the national level, research by economic historians has emphasized the costs associated with World War II (Higgs 2006; Rockoff 2012) and improvements in the technology and public infrastructure underlying the United States economy in the 1930s relative to the 1940s (Field 2011). For local economies, work by Fishback and Cullen (2013) has studied the short-run impact of the war on measures of consumption and income, with a focus on the costs of adjusting from a wartime economy.

In the second part of this chapter we extend recent work on the impact of World War II on local economies. We use a simple spatial equilibrium model to understand the link between mobilization for World War II and long-run changes at the county level. Specifically, we use data for all United States counties to estimate the long-run impact of the war-related spending on income per capita, population, and median housing values per decade from 1960 to 2010. Using the estimated regression coefficients and the parameters for the model to examine the impact of the war on the amenities, productivity, and housing supply.

World War II and Macroeconomic Indicators

World War II brought sweeping changes to the United States economy. Those who lived through the war and its aftermath witnessed tectonic shifts in the size and control of the national government at home as well as a transformation of the ability and willingness to project American power beyond its territorial boundaries. As a result, national institutions governing economy and society played an increasingly important role in the daily lives of Americans.

Together with improvements in macroeconomic indicators over the war and postwar periods, economists came to view the changes brought on by mobilization for World War II as a key determinant of economic growth in the second half of the twentieth century. During the war, civilian unemployment fell from 9.5 % in 1940 to below 2 % in 1943, 1944, and 1945, and income per person reached a wartime peak in 1944 that was 67 % higher than the 1940 level. By 1945, gross domestic product had increased 70 % and per capita consumption was up nearly 25 %.

The connection between mobilization and gains in these aggregate indicators is typically done by applying a fiscal policy multiplier to military spending that subsequently spilled over into the civilian economy. Yet, by the end of the 1941 the economy had nearly returned to its long-run growth path. Real GDP per capita was 18 % above its 1929 peak, implying an annual rate of growth of 1.5 % over the intervening years. At the same time, the employed share of the population had reached 39.4 %, or 1.4 % above the pre-Depression peak, and the unemployment rate had fallen to 6 %.[1]

Once the United States entered the War, the American economy took on many of the trappings of a command economy. At the height of the war over 40 % of GDP was devoted to the production of military goods, including 17 million rifles and pistols, over 80,000 tanks, 41 billion rounds of ammunition, 4 million artillery shells, 75,000 vessels, and nearly 300,000 planes. In many cases, the production

[1] Real GDP is series Ca11 (p. 3–25), the information on employment and unemployment came from series Ba470, Ba474, Ba475, and Ba777 (pp. 2–82, 2–83), population is series Aa7, pp. 1–28, 1–29. These data are from Carter et al. (2006). The labor markets were still influenced by the presence of emergency work relief agencies like the Works Progress Administration, which provided work for poor families and paid for poor families at roughly half pay. The employment and unemployment rate figures above treat such workers as employed. If they are treated as unemployed, the employed as a share of population is 37.7 %, 0.3 % lower than in 1929 and the unemployment rate is 9.9 %.

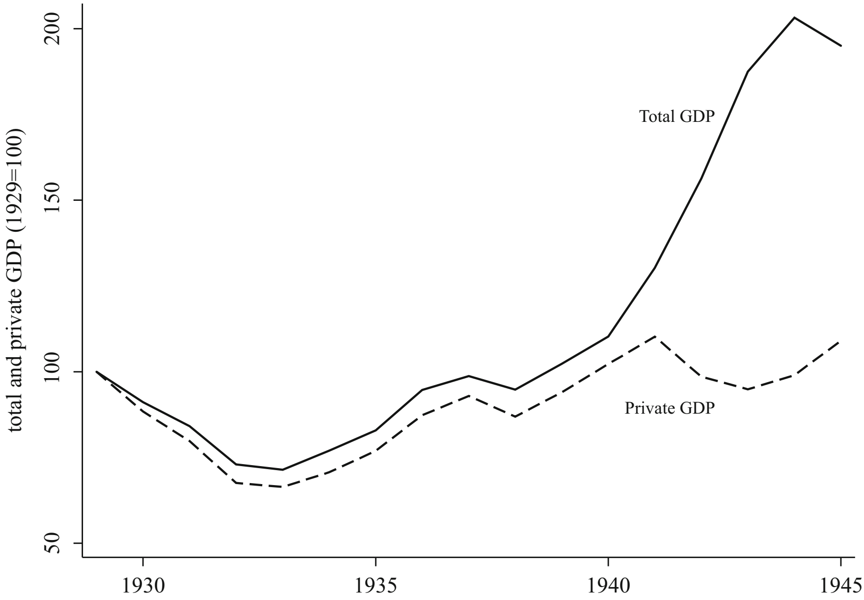


Fig. 1 Total and private gross domestic product, 1929–1945 (Notes: Authors' calculations based on Council of Economic Advisors (1995, p. 406))

of military goods substituted for civilian goods. For example, tanks replaced automobiles, anti-aircraft mounts replaced electric washing machines, and nearly all rubber went to the war effort. In the aggregate this shows up in the divergence in total and private GDP during the war years shown in Fig. 1. Ultimately, this led to rationing on the home front.

In addition, the war years were not accompanied by increased total factor productivity Field (2008). Indeed, compared with the Great Depression and the postwar years, the efficiency of the economy did not improve. Instead, nearly all research and development was centered on military activities. There were some benefits of these activities that spilled over to civilian uses–such as in the development of penicillin, improved techniques for the production of ships and planes, microwave technology, and advances in electronics. However, many other innovation were specific and could not easily be adapted to civilian uses. Thus, the dramatic increase in output between 1940 and 1945 was achieved through the allocation of more inputs (e.g., labor, capital, and natural resources) to war production.

Fighting the war required the conscription of ten million people into the armed forces and another six million volunteers shown in Fig. 2, that is, up to 18 % of the combined civilian and military labor force in 1945. In the absence of wages high enough to draw men into military service, the draft was necessary to meet manpower requirements from the armed forces. A private first class with no dependents could

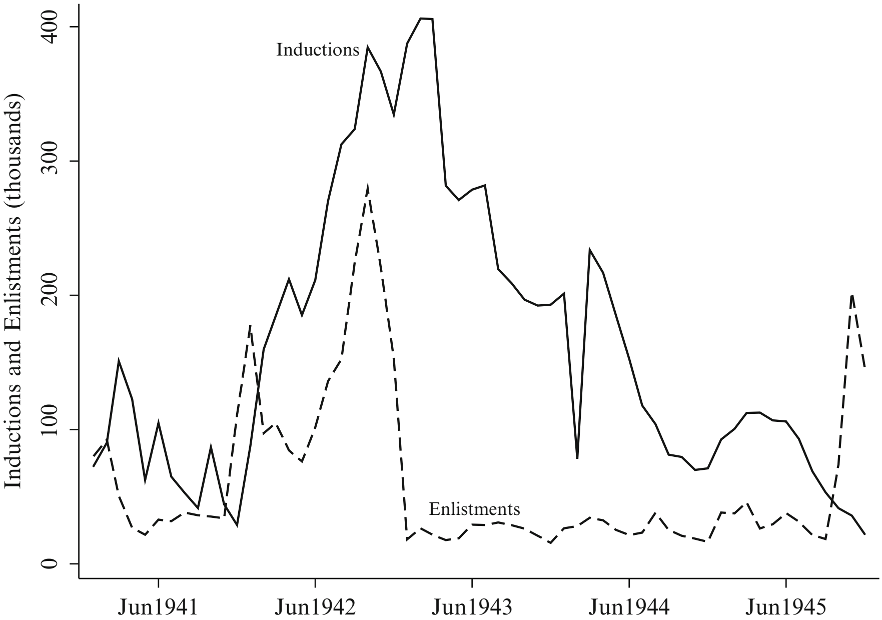


Fig. 2 Monthly enlistments and inductions, 1941–1945 (Notes: The number of inductions (*solid line*) and enlistments (*dashed line*) by month from January 1941 to December 1945. The sharp decrease in inductions in February 1944 is due to initiation of a pre-induction physical examination. The data are drawn from the U.S. Selective Service System (1948, pp. 32–33))

expect to earn $58 per month compared with $184 per month for an unskilled worker in manufacturing in 1943. Using a simple dynamic general equilibrium model, Siu (2008) finds that the welfare loss associated with substituting conscription for fiscal policy (i.e., increasing expenditures on salaries for servicemen) up to 2 % of aggregate annual consumption. This calculation excludes the costs associated with the death of 400,000 soldiers and injuries to additional 670,000.

Hours per week rose from 38 in 1940 to 45 in 1944 in manufacturing. Night shifts became more common and disabling manufacturing injury rates rose. Moreover, although official figures for real consumption show no change between 1941 and 1944, these do not account for the decrease in product quality, the costs of obtaining rationed goods, and the absence of a wide variety of goods from the marketplace. Production halted for cars, tires, and many other household goods. The Office of Price Administration set price ceilings on a wide variety of goods and limited access to meat, gasoline, fuel oil, kerosene, nylon, silk, shoes, sugar, coffee, processed foods, cheese, and milk. Adjusting the consumption figure to reflect these additional costs, consumption per person was likely lower.

To pay for the war, the federal government greatly expanded the number of taxpayers. The income level at which households were required to begin paying income taxes was cut from $2000 to $600 for individuals and from $5000 to $600 for a family of four. Between 1940 and 1945 the number of federal income taxpayers rose from 7.1 million to over 42 million, nearly a sixfold increase. In addition, at the top end of the income distribution, the average tax rate was 90 %. As a result, federal tax collections rose from 6.5 % of GDP in 1940 to 20 % in 1945.

In the end, the use of fiscal policy during World War II offers few lessons for its application in peacetime. Early work, for example due to Vernon (1994), argued that most of the recovery from the trough of the Great Depression can be explained by war-related fiscal policy. This work was based on a fiscal multiplier of 1.6. For the World War II period, Barro (1981) estimated an income multiplier of 0.6, which suggests that an additional dollar World War II spending led to less than an additional dollar of income. In contrast, immediately prior to outbreak of World War II, Gordon and Krenn (2010) estimated a multiplier of 1.8 in an environment that included high unemployment and multiplier of 0.9 when output and potential output moved closer together.

Mobilization and the Development of Local Economies

Planning for mobilization in World War II was informed by the experience during World War I. Unprepared for the technological and logistical sophistication of modern war, civilian and military leaders in the 1920s and 1930s sought to address the coordination problems and organizational failures that plagued mobilization. The War Industries Board, which was disbanded in January 1919, was revived as the National Recovery Administration in 1933. In addition, other World War I agencies–including the Grain Corporation, the Emergency Fleet Corporation, the Fuel Administration–reemerged as the specter of war increased.

The 1920 National Defense Act contained new directives to ensure "the adequate provision of the mobilization of material and industrial organizations essential to wartime needs." The Act emphasized the need for extensive planning to coordinate the many economic dimensions of mobilization for modern warfare. This brought attention to planning for procurement and, eventually, economic mobilization among civilian administrators and plans by the armed forces began to recognize the need to balance military and civilian requirements. Prewar planning culminated in several Industrial Mobilization Plans, the last in 1939, which started to indicate the stark regional differences in planning for mobilization.

Following events in Europe between April and June of 1940, American public opinion toward rearmament began to soften. President Roosevelt's fireside chat of May 26, 1940, laid out a broad mobilization program "calling upon the resources, the efficiency and the ingenuity of the American manufacturers of war material of all kinds." Two days later, the National Defense Advisory Commission was established to oversee and coordinate industrial production, and legislation passed in early summer and fall re-authorized the use of cost-plus contracts in place of competitive bidding, allowed for accelerated depreciation of war-related capital expenditures, and removed the excess profit provisions on government contracts.

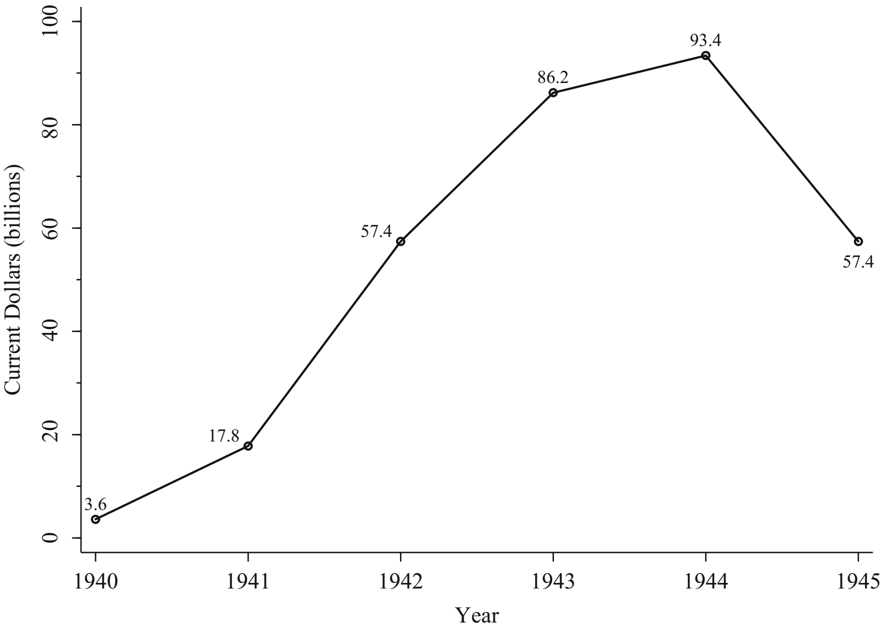


Fig. 3 World War II spending, 1940–1945 (Notes: Estimates reported in this figure were originally compiled by the War Production Board and are drawn from Smith (1959, p. 7))

Spending on supply contracts and direct (or indirect) investment account for the vast majority of government expenditures and provide the basis for recent empirical work in economic history. Total war spending, shown in Fig. 3, increased to \$93.4 billion in 1944 from \$3.6 billion in 1940. Roughly half of the facilities producing for the war were located in the most industrialized areas of the Northeast and Upper Midwest to make use of existing facilities for producing the large quantity of goods required by the military. Still, for reasons including patronage, security, congestion, and weather, as well as (more importantly) the availability of labor, raw materials, and land, a large share of spending went to counties in the South and West.

The domestic boom in military production spurred migration to cities. In turn, the federal government took steps to ameliorate the negative consequences. For example, in the fall of 1940, the Defense Housing and Community Facilities and Services Act was passed and included provisions to provide funding for housing and other public services in areas deemed vital for the national defense. In addition, under the Servicemen's Dependents Allowance Act of 1942, the emergency maternity program, and the infant care program, families of service personnel were provided with small payments and health-care benefits. Finally, to combat rising rents, the Emergency Price Control Act of 1942 empowered the Office of Price Administration to lower rents in the some areas to their prewar levels.

The relative size of the mobilization compared with some regional economies led early scholars of the period to emphasize the war's role in increased industrial activity in the states along the Pacific Coast and in the South. According to Nash (1990), California, and the West more generally, primarily provided natural resources to the industrial centers in the Midwest and Northeast prior to the war. However, mobilization led to new infrastructure and provided the impetus to learn new production techniques, in particular through the investment in new capital and supply contracts for more advanced goods. In contrast, Rhode (2003) argues that mobilization simply reinforced the growth in manufacturing that was already under way. Perhaps the war facilitated more rapid industrialization, but the level of development achieved in the immediate aftermath of World War II likely would have been achieved eventually.

For the American South, the potential for the war to shock a low-wage and labor-intensive economy was perhaps greater. In particular, mobilization may have helped attract manufacturing to the region: "[A] wave of new plants came to the South because of natural resources, climate, and a labor pool attractive for its size if not always for its skills .... During the war, they helped train a managerial group whose entrepreneurial skills were a continuing asset to the South and acquainted many of the rural poor with an alternative way of life" (White 1980). Lewis (2007) argues that this did not occur. Rather, Southern manufacturing lacked the strong linkages between firms typically needed to generate for spillovers from government spending that benefit private producers. Industrial output did increase during the war years, but Jaworski (2015) provides evidence that this did not lead to the region's postwar industrialization.

Overall, the evidence suggests that mobilization for World War II was not responsible for recovery from the Great Depression nor did it lead to the industrialization of the regional economies. Fishback and Cullen (2013) provide direct evidence for the limited short-run effect of the war (i.e., through the late 1950s) on economic activity in the form of retail sales. This finding contrasts sharply with the results for the impact of government spending on public works and relief during the New Deal: an additional dollar of public works and relief spending per capita between 1933 and 1939 raised retail sales per capita by about 40–50 cents in 1939 (Fishback et al. 2005). There are several reasons for this difference. First, New Deal spending was done in an environment with unemployment as high as 25 %, so that many unemployed resources could be put to use before crowding out occurred. Second, New Deal spending mostly went to public infrastructure projects that were complementary to and not substitutes for private production (see also Field 2011). Third, the New Deal did not reach the levels of World War II spending, never accounting for more than 8 % of national income. Finally, the capital used during the war was not easily converted to civilian uses.

Fishback and Cullen (2013) find that the strongest effect of World War II on local economies was on population growth. This is consistent with evidence provided by Blanchard and Katz (1992) and a wide class of spatial equilibrium models in which changes in population are the main way that local economies adjust to economic shocks. The remainder of this chapter formalizes a simple spatial equilibrium model to study the long-run effect of World War II on local economies.

The Impact of World War II in Spatial Equilibrium

As a framework for the empirical analysis presented below, we draw on the dynamic version of the Rosen–Roback spatial equilibrium model presented in Glaeser (2008). The model provides expressions for the equilibrium change in population, income per capita, and housing prices that can be used to assess the contribution of World War II spending to growth in productivity, amenities, and the supply of housing. In the model, each local economy has three sectors: firms produce a nationally traded good, individuals choose location and consumption, and builders provide housing. Consumers maximize a Cobb–Douglas utility function,

$$U = \theta_t C_t^{1-\alpha} H_t^{\alpha}, \tag{1}$$

by choosing the quantity of tradable goods, C_t, and consumption of non-traded housing, H_t.

We can simplify the consumer problem by replacing $C_t = W_t R_t H_t$, where consumers use all of their income W_t to purchase either housing at the rental price R_t or the tradable good (with a fixed price of 1). This specification of the problem yields an indirect utility function

$$U = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\theta_t W_t P_t^{-\alpha} \tag{2}$$

The share of household income spent on housing is α. The spatial equilibrium assumption says that consumers must be indifferent across counties, which is equivalent to setting indirect utility equal to a reservation utility, $\underline{U}$. In this way, we allow workers to move across locations (i.e., counties) following a shock to the local economy in the form of World War II spending.

In the production sector, a firm in time t produces final output using the following constant returns to scale technology

$$Q_t = A_t N_t^{\beta} K_t^{\gamma} Z_t^{1-\beta-\gamma} \tag{3}$$

where A_t is county-specific total factor productivity and $N_t, K_t,$ and Z_t are labor, traded capital, and non-traded capital, respectively. Traded capital has a price of 1 in all locations and there is a fixed supply of non-traded capital $\overline{Z}$. Firm profit maximization yields first-order conditions for input choices that can be used to solve for labor demand. To pin down the size of the traded good sectors, we allow new firms to enter each location until firm profits are equal to zero.

The construction sector is dynamic and has developers choosing in each period whether to build and sell a house in order to maximize $\frac{P_t - c_t}{(1+r)^t}$, where c_t is the period t cost of construction. The solution to this problem is a cutoff rule specifying in which period to develop a piece of land. In particular, the decision to build is based on growth in construction costs, the interest rate, and the endogenously determined price of housing. This gives an expression for total housing supply as a function of price, where ρ is the elasticity of housing supply with respect to price.

When each sector is in equilibrium, a system of three equations with three unknowns is generated that can be solved for the equilibrium growth in population (N_t), income (W_t), and housing prices (P_t). These expressions are given below

$$\log\left(\frac{N_{t+1}}{N_t}\right) = \frac{(1+\rho-\alpha)\log(1+g_A)+(1-\gamma)\left((\rho+1)\log\left(\frac{1+g_\theta}{1+g_U}\right)-\alpha\rho\log(1+g_H)\right)}{(1-\beta-\gamma)(\rho+1)+\alpha\beta} \tag{4}$$

$$\log\left(\frac{W_{t+1}}{W_t}\right) = \frac{\alpha\log(1+g_A)+(1-\beta-\gamma)\left(\alpha\rho\log(1+g_H)-(\rho+1)\log\left(\frac{1+g_\theta}{1+g_U}\right)\right)}{(1-\beta-\gamma)(\rho+1)+\alpha\beta} \tag{5}$$

$$\log\left(\frac{P_{t+1}}{P_t}\right) = \frac{\log(1+g_A)+\beta\log\left(\frac{1+g_\theta}{1+g_U}\right)+\rho(1-\beta-\gamma)\log(1+g_H)}{(1-\beta-\gamma)(\rho+1)+\alpha\beta} \tag{6}$$

where g_A is the growth rate in productivity, g_θ is the growth rate in the amenities of living in the community, and g_H is the growth rate of construction costs, and g_U is the growth rate in the utility.

Equations (4), (5) and (6) provide the theoretical basis for empirical work that examines the effect of World War II spending on local economic development. In particular, we can write the expressions for the growth in productivity, amenities, and construction costs as

$$\log(1+g_A) = \kappa_A + \lambda_A WWII_{ct} + \mu_A \tag{7}$$

$$\log\left(\frac{1+g_\theta}{1+g_U}\right) = \kappa_\theta + \lambda_\theta WWII_{ct} + \mu_\theta \tag{8}$$

$$\log(1+g_H) = \kappa_H + WWII_H\, X_{ct} + \mu_H \tag{9}$$

where $WWII_{ct}$ is a vector of war spending, $\kappa_A, \kappa_\theta,$ and κ_H are constants, $\lambda_A, \lambda_\theta,$ and λ_H are coefficients to be estimated, and $\mu_A, \mu_\theta,$ and μ_H are error terms. Since we do not directly observe measures of amenities, productivity, or housing supply, we use measures of population (N_t), income per capita (W_t), or housing prices (P_t), together with Eqs. (3), (4), (5), (6), (7), (8), and (9) to estimate the impact of different types of war-related spending on productivity, amenities, and housing.

We use data on population, income per capita, and median housing values in 1940 and from 1960 to 2010 to estimate the effect of each two type of war spending:

$$\log\left(\frac{N_{it}}{N_{i,1940}}\right) = \pi_{1,t}^{N} \log\left(\text{contracts}_{i,1940-45}\right) + \pi_{1,t}^{N} \log\left(\text{capital}_{i,1940-45}\right) + X_{it}\delta^{N} + \varepsilon_{it}^{N} \tag{10}$$

$$\log\left(\frac{W_{it}}{W_{i,1940}}\right) = \pi_{1,t}^{W} \log\left(\text{contracts}_{i,1940-45}\right) + \pi_{1,t}^{W} \log\left(\text{capital}_{i,1940-45}\right) + X_{it}\delta^{W} + \varepsilon_{it}^{W} \tag{11}$$

$$\log\left(\frac{P_{it}}{P_{i,1940}}\right) = \pi_{1,t}^{P} \log\left(\text{contracts}_{i,1940-45}\right) + \pi_{1,t}^{P} \log\left(\text{capital}_{i,1940-45}\right) + X_{it}\delta^{P} + \varepsilon_{it}^{P} \tag{12}$$

Using the log difference between the outcome in year t and 1940 allows us to address concerns that our estimates reflect the prewar characteristics of counties. The coefficients of interest are those on per capita supply contracts $(\text{contracts}_{c,1940-45})$ and capital investment $\left(\text{capital}_{c,1940-45}\right)$, including facilities and equipment investment, between 1940 and 1945 in county i. We estimate a coefficient for each type of spending and each postwar decade to capture changes over time in local economic development related to World War II. In addition, in X_{it}, we include additional controls for state-year fixed effects to focus on estimates that compare counties that share the same state policy environment as well as lagged 1940 values of each outcome variable (interacted with year fixed effects) to control for mean reversion in population, income per capita, or housing prices.

After reporting the results for each type of war-related spending on population, per capita income, and housing prices, we then use the estimated coefficients to calculate the impact of World War II spending (either supply contracts and capital investment). The effects for each are shown by the following equations:

$$\lambda_{Aj} = (1-\beta-\gamma)\,\widehat{\pi}_j^{N} + (1-\gamma)\,\widehat{\pi}_j^{W} \tag{13}$$

$$\lambda_{\theta j} = \alpha\widehat{\pi}_j^{P} - \widehat{\pi}_j^{W} \tag{14}$$

$$\lambda_{Hj} = \widehat{\pi}_j^{P} + \frac{1}{\rho}\left(\widehat{\pi}_j^{P} - \widehat{\pi}_j^{N} - \widehat{\pi}_j^{W}\right) \tag{15}$$

The term λ_{Aj} says that productivity growth is the weighted average of the effect on population and income per capita with weights from the Cobb–Douglas production function. The second term, $\lambda_{\theta j}$, says that amenity growth due to each type of war spending is the difference between the war spending effect on housing prices times the share of income spent on housing minus the effect of war spending on wages. The third term, λ_{Hj}, shows that impact of the war spending on housing is a function of its on population, housing prices, and wages and the price elasticity of the supply of housing.

Data

To estimate the coefficients associated the Rosen–Roback framework we make use of data on income per capita, total population, and median housing values at the county level from 1940 and 1960 to 2010. The data on income per capita in 1940 are drawn from the complete count of the 1940 census Ruggles et al. (2015) and population and median house values from Haines (2010). Income per capita, population, and median housing values for 1960–2010 are drawn from Haines (2010) and the Bureau of Economic Analysis (2015). Our sample consists of the 3051 counties for which data is available in each year.

For World War II spending at the county level, we draw on four variables available in Haines (2010) that capture separate aspects of the mobilization effort between 1940 and 1945: supply contracts, combat equipment contracts, investment in industrial facilities, and investment in military facilities. Supply contracts were prime contracts awarded by the War Department, Navy, Maritime Commission, Treasury Procurement Division, and foreign purchasing greater than $50,000, reported to the War Production Board, and assigned to a county where principal production would take place. The total value of the contracts are net any reductions or cancellations. Contracts for combat equipment were tabulated separately and included aircraft, ships, and ordnance material. Our first measure of county-level spending on World War II is the sum of spending on supply contracts and combat equipment contracts.

Facilities projects included improvements to existing facilities or the construction of new facilities financed by the Army, Navy, Maritime Commission, Defense Plant Corporation, Reconstruction Finance Corporation, or British Empire governments, and valued at more than $25,000. Projects for industrial and military facilities were recorded separately. Spending on industrial facilities went to improvements or new construction of establishments producing war material. For military facilities, spending went to temporary barracks, airports, and other military installations and was financed by the armed forces. The second measure of World War II spending is the sum of investment in industrial facilities and investment in military facilities.

Figure 4 shows the distribution of per capita spending on supply contracts (Panel A) and capital investment (Panel B). Across all counties average spending per capita was $408 on supply contracts and $174 on capital investment. However, close to one half of counties received no spending on supply contracts and over two-thirds received no capital investment. Overall, spending was concentrated in the heavily industrialized areas and the top 20 counties accounted for nearly 40 % of total spending, particularly in the counties containing Detroit, Los Angeles, Chicago, New York City, and Buffalo.

Finally, to estimate the relative importance of growth in productivity, amenities, and housing supply using Eqs. (5), (6), and (7), we need to assume values for α, β, γ, and ρ, which we take from Glaeser and Tobio (2008). In particular, the share of household income spent on housing α is set to 0.3, β and γ correspond to the share

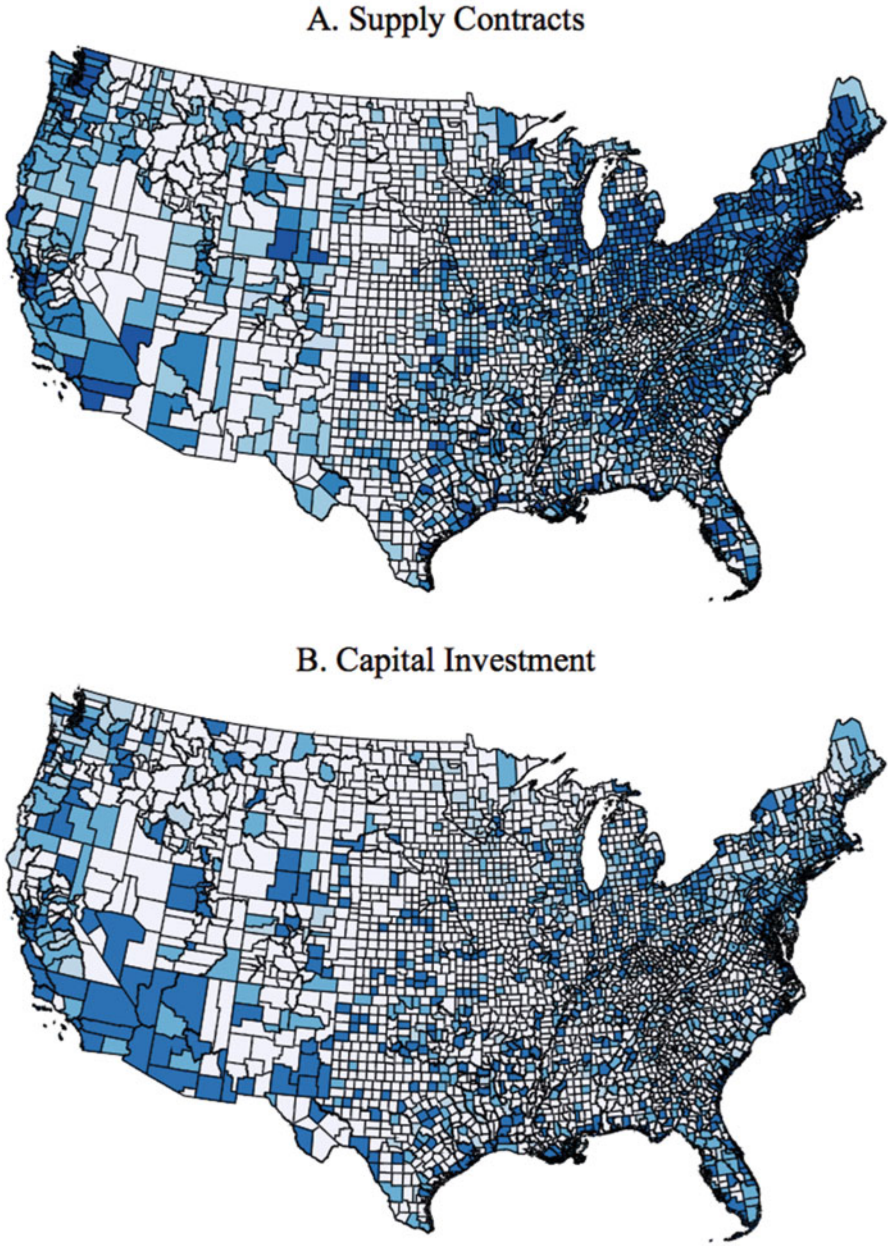


Fig. 4 Location of World War II spending, 1940–1945 (Notes: The value of supply contracts (*Panel A*) and capital investment (*Panel B*) expenditures per capita between 1940 and 1945. *Darker shades of blue* indicate more spending per capita. Data supply contracts and capital investment are drawn from Haines (2010)

of labor and traded capital in input costs and are set to 0.6 and 0.3, respectively, and ρ is the housing supply elasticity, which is set to either 1.5 or 3.0.

Empirical Results

We now turn to an assessment of the influence of World War II spending per capita on local economic development. First, the results coefficients from estimating Eq. (4) are shown in Fig. 5 together with the 95 % confidence interval based on standard errors clustered at the state level to allow for correlation across counties within the same state over time. Results for income per capita, total population, and median house value are shown in each panel. In each case, the impact of supply contracts and capital investment on the outcome variables is plotted separately.

In Panel A, the results for the impact of per capita war spending on income per capita after 1960 are consistent with the results from earlier studies that proxy for income with retail sales per capita. For both types of war spending the estimated coefficients are close to zero. In 2000, a 10 % increase in spending on war-related supply contracts increased income per capita by 0.06 %, while the same increase in capital investment increased income per capita by 0.07 %. The effect of capital investment is statistically significantly different from zero over much of the postwar period, but for both types of spending the effect is economically very small.

Panel B shows that the main adjustment in local economies was due to changes in total population. For both types of spending the effect is statistically significant and tends to increase between 1960 and 2010. In 1960, a 10 % increase in World War II spending was associated with an increase in population by 0.28 % for supply contracts and 0.47 % for capital investment; by 2010, these magnitudes had increased to 0.45 and 0.64 %, respectively. These effects represent a substantial reallocation of population to centers of war production and may reflect the war's role in facilitating a major structural change in employment away from agriculture toward manufacturing and services.

Finally, Panel C shows the effect of each type of war spending on median house values. The housing sector is an important aspect of the economy when seeking to understand the ability of local economies to adjust to large economic shocks. In particular, gains in income per capita in a county may be partially (or entirely) offset by increases in housing prices. In addition, improvements in amenities or local public goods (i.e., better public health and schooling) may be capitalized into the value of the housing stock. The estimates in Panel C show that an additional 10 % in war supply contract spending in the 1940s was associated with roughly 0.12 % higher median housing prices in every census year between 1960 and 2010. In contrast, the impact of war capital spending fell over time; 10 % higher spending was associated with 0.19 higher median housing values in 1960 and the effect decline to 0.10 % by the 2000s.

Overall, the impact of war spending on housing prices was larger than for income per capita throughout the postwar period, but this did not deter the flow of people

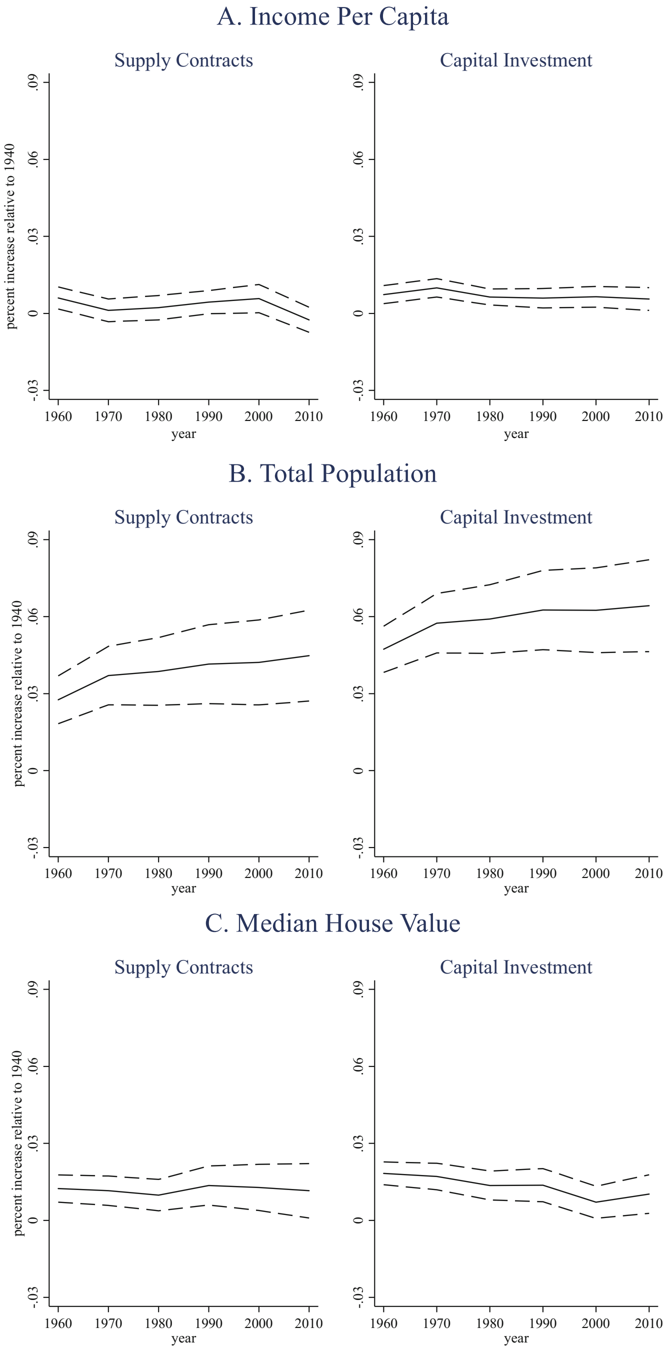


Fig. 5 Impact of World War II on income, population, and housing (Notes: The figure plots the coefficient for the impact of supply contracts and capital investment on the growth of income per capita (*Panel A*), total population (*Panel B*), and median housing value between 1940 and a given year from Eqs. (10), (11), and (12), respectively)

Table 1 Impact of supply contracts on amenities, productivity, and housing supply

	1960	1970	1980	1990	2000	2010
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: all counties						
Amenities	0.0070	0.0045	0.0054	0.0072	0.0082	0.0027
Productivity	−0.0023	0.0023	0.0007	−0.0003	−0.0019	0.0060
Housing supply w/$\rho = 1.5$	−0.0017	−0.0062	−0.0108	−0.0079	−0.0105	−0.0090
Housing supply w/$\rho = 3.0$	0.0053	0.0027	−0.0005	0.0028	0.0011	0.0013
Panel B: northeast						
Amenities	0.0019	0.0007	−0.0020	−0.0044	−0.0067	−0.0074
Productivity	0.0009	−0.0029	−0.0018	−0.0009	0.0002	−0.0024
Housing supply w/$\rho = 1.5$	−0.0102	−0.0327	−0.0411	−0.0498	−0.0499	−0.0665
Housing supply w/$\rho = 3.0$	−0.0041	−0.0233	−0.0313	−0.0394	−0.0414	−0.0551
Panel C: midwest						
Amenities	0.0055	0.0059	0.0042	0.0081	0.0072	0.0048
Productivity	−0.0002	0.0010	0.0019	−0.0009	−0.0018	0.0030
Housing supply w/$\rho = 1.5$	0.0014	−0.0055	−0.0136	−0.0075	−0.0210	−0.0186
Housing supply w/$\rho = 3.0$	0.0072	0.0038	−0.0030	0.0037	−0.0071	−0.0050
Panel C: south						
Amenities	0.0111	0.0069	0.0066	0.0068	0.0059	0.0015
Productivity	−0.0073	−0.0020	−0.0022	−0.0021	−0.0008	0.0053
Housing supply w/$\rho = 1.5$	−0.0021	−0.0100	−0.0141	−0.0160	−0.0143	−0.0132
Housing supply w/$\rho = 3.0$	0.0063	−0.0001	−0.0034	−0.0043	−0.0033	−0.0034
Panel D: west						
Amenities	0.0025	0.0040	0.0039	0.0090	0.0140	0.0050
Productivity	−0.0008	−0.0001	0.0039	−0.0034	−0.0093	0.0036
Housing supply w/$\rho = 1.5$	−0.0093	−0.0040	0.0144	0.0217	0.0279	0.0399
Housing supply w/$\rho = 3.0$	−0.0039	0.0019	0.0170	0.0228	0.0286	0.0353

The table shows the effect of supply contracts on amenities, productivity, and housing supply growth in each postwar decade relative to 1940. *Panel A* shows the results for all U.S. counties together and the *Panels B through E* show the results for each census region

to counties with substantial World War II spending. Such a finding suggests that long-run changes in local economies due to World War II did not reflect productivity gains that raised local wages. Instead, these findings point toward benefits associated with mobilization that were primarily non-pecuniary. And this is consistent with the recent literature that the direct benefits, i.e., in terms of consumption or income, of mobilization were small.

Using Eqs. (13), (14), and (15) we can calculate the effects of World War II spending on productivity, amenities, and housing supply. Tables 1 and 2 present the results of this exercise. The columns show the results for each postwar decade relative to 1940. In each table, Panel A uses estimates pooling all counties together, while the remaining panels show the results for each census region (i.e., Northeast, Midwest, South, and West) separately.

Table 2 Impact of capital investment on amenities, productivity, and housing supply

	1960	1970	1980	1990	2000	2010
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: all counties						
Amenities	0.0098	0.0127	0.0104	0.0104	0.0108	0.0103
Productivity	−0.0018	−0.0048	−0.0023	−0.0018	−0.0044	−0.0025
Housing supply w/$\rho = 1.5$	−0.0059	−0.0164	−0.0210	−0.0228	−0.0342	−0.0295
Housing supply w/$\rho = 3.0$	0.0062	0.0003	−0.0037	−0.0045	−0.0136	−0.0096
Panel B: northeast						
Amenities	0.0028	0.0049	0.0054	0.0096	0.0094	0.0097
Productivity	0.0030	0.0002	−0.0019	−0.0087	−0.0078	−0.0074
Housing supply w/$\rho = 1.5$	0.0036	−0.0103	−0.0158	−0.0256	−0.0243	−0.0208
Housing supply w/$\rho = 3.0$	0.0078	−0.0012	−0.0057	−0.0122	−0.0109	−0.0079
Panel C: midwest						
Amenities	0.0052	0.0054	0.0043	0.0042	0.0044	0.0041
Productivity	−0.0008	−0.0013	−0.0008	0.0003	−0.0008	0.0010
Housing supply w/$\rho = 1.5$	−0.0093	−0.0169	−0.0229	−0.0208	−0.0264	−0.0210
Housing supply w/$\rho = 3.0$	−0.0008	−0.0060	−0.0110	−0.0089	−0.0132	−0.0087
Panel C: south						
Amenities	0.0112	0.0160	0.0127	0.0127	0.0141	0.0141
Productivity	−0.0029	−0.0069	−0.0022	−0.0009	−0.0046	−0.0030
Housing supply w/$\rho = 1.5$	−0.0015	−0.0103	−0.0130	−0.0146	−0.0261	−0.0214
Housing supply w/$\rho = 3.0$	0.0100	0.0068	0.0044	0.0043	−0.0044	−0.0002
Panel D: west						
Amenities	0.0100	0.0103	0.0093	0.0086	0.0088	0.0071
Productivity	−0.0008	−0.0003	0.0007	0.0016	−0.0005	0.0026
Housing supply w/$\rho = 1.5$	−0.0149	−0.0221	−0.0276	−0.0320	−0.0468	−0.0462
Housing supply w/$\rho = 3.0$	0.0010	−0.0031	−0.0073	−0.0105	−0.0217	−0.0211

The table shows the effect of capital investment on amenities, productivity, and housing supply growth in each postwar decade relative to 1940. *Panel A* shows the results for all U.S. counties together and the *Panels B through E* show the results for each census region

For all United States counties the results suggest a consistent role of improved economic welfare due to amenity growth in locations with supply contracts and capital investment related to World War II. For supply contracts Table 1 shows that an additional 10 % of spending increased amenity growth relative to 1940 by 0.03–0.08 % between 1960 and 2010. For capital investment in Table 2 the role in amenity growth in promoting economic welfare is larger: 0.09–0.13 for an additional 10 % of spending. In contrast, the contribution of productivity growth associated with war-related spending to economic welfare tended to be small in magnitude and negative for all years for war capital spending and for half of the years for contract spending.

The rise in welfare due to the impact of war contract spending on housing supply is small. We experimented with different elasticities for the responsiveness of housing supply to prices of 1.5 and 3. When the elasticity assumption is changed from 1.5 to a more elastic 3, the contribution of the war contract impact on housing supply to economic welfare switches from negative to positive. However, the effect is small in both cases.

To some extent, the results for all counties may mask regional differences in the effect of World War II due to the level of industrial development prior to 1940 or different underlying regional trends. Panels B through E of Tables 1 and 2 show the decomposition of changes in economic welfare in each census region separately. The results reveal that regional differences may be important for understanding the long-run impact of World War II on local economic development.

Amenity growth related to supply contracts was small and positive in the Northeast in 1960 and 1970 and slightly negative thereafter. This is mitigated by the positive effect of capital investment on amenities in the Northeast. Taken together these results suggest that for counties in the Northeast where war spending came only in the form of supply contracts and therefore was temporary, amenities deteriorated. This effect was less stark when supply contracts were accompanied by capital investment. The difference in the effect of the two types of spending on amenities persisted until 2010, which suggests that even several decades after the war-related capital was installed people were still attracted to the places it had been.

In the Midwest, South, and West the effect of supply contracts and capital investment on amenities reinforced one another. This result may reflect the continued role these regions played in the growth of the military-industrial complex during the Cold War, in particular, through defense contracts as well as employment related to military bases (Markusen et al. 1991). Local politicians and businessmen worked hard to attract contracts and investment to communities, in particular, in the South and West (Combes 2001; Koistinen 2004). Contracts and investments in these regions are not associated with significant gains in productivity. However, they may have helped to attract additional government resources as well as human capital that were eventually reflected in the value of local amenities.

For housing supply, the effect of supply contracts and capital investment is similar across regions, although the choice of the housing supply elasticity plays a role in assessing its importance relative to amenities and productivity. In general, the effect of World War II on housing supply starts close to zero in 1960 and becomes increasingly negative until 2010. The one exception is for the West, which experienced an expansion of housing supply over the postwar period.

Overall, the results of this exercise show that changes in population represent the most consistent and persistent response to mobilization for World War II. Even several decades later population continued to grow in places where value of per capita spending on supply contracts and capital investment was large during World War II. These changes were not accompanied by significant gains in per capita income or rapid growth in the value of housing. This suggests that workers were taking advantage of amenities that capitalized—or, at least, were correlated with—the value of war-related spending. These amenities may have come in the form of proximity to defense-related industries in the postwar period that increased employment opportunities, access to local public goods, or proximity to high levels of human capital.

Conclusions

The size and extent of mobilization for World War II was unprecedented in United States history. Economic historians of this period have recently focused on revising the traditional narrative that attributed much of the growth during the postwar "golden age" to technological changes brought on by World War II. Field (2011) emphasized the importance of total factor productivity growth prior to 1940, while Higgs (1987) and Rockoff (2012) emphasized the costs associated with mobilization and demobilization as well as the persistence in institutions (e.g., Selective Service and the militaryindustrial complex). More recent work has examined the short-run impact on consumption (Fishback and Cullen 2013) and industrialization (Rhode 2003; Bateman et al. 2009; Jaworski 2015) at the local or regional level.

In this chapter, we use a simple spatial equilibrium model to assess the long-run impact of mobilization for World War II on local economic development. The empirical results show that the main changes in local economies were due to the reallocation of population toward counties that received more per capita war spending. In the context of limited growth in income per capita and median house values, the interpretation provided by the model suggests that mobilization for World War II was mostly correlated with an increase in the value of local amenities. In practice, the growth in amenities is likely related to the proximity to defense-related industries and associated non-wage benefits.

The results point toward new avenues for research into the impact of World War II. In particular, the findings suggest that scholars might focus less on drawing direct analogies from fiscal multipliers during wartime in discussing multipliers during peacetime. As an alternative, they might focus more on the political economy of benefits and costs of wartime spending at the state and local level. Importantly, this includes the development of the military-industrial complex through the interactions of local politicians and businessmen with the national government, the transfer of technology between public and private sector activities, and the determinants of the distribution of federal defense-related spending across locations and firms.

References

Barro, Robert J. 1981. Output effects of government purchases. *Journal of Political Economy* 89(6): 1086–1121.

Bateman, Fred, Jamie Ros, and Jason E. Taylor. 2009. Did new deal and World War II public capital investments facilitate a 'Big Push' in the American South? *Journal of Institutional and Theoretical Economics* 165(2): 307–341.

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. *Regional evolutions*. Brookings Papers on Economic Activity, pp. 1–75.

Carter, Susan, Sigmund Carter, Scott, Haines, Michael, Olmstead, Alan, Sutch, Richard, and Wright, Gavin. 2006. *Historical statistics of the United States: Earliest times to the present, millennial edition.* New York: Cambridge University Press.

Combes, Richard S. 2001. Aircraft manufacturing in Georgia: A case study of federal industrial investment. In *The second wave: Southern industrialization from the 1940s to the 1970s*, ed. Philip Scranton. Athens: University of Georgia Press.

Council of Economic Advisors. 1995. Economic report of the president: Transmitted to the Congress February 1995. United States Government Printing.

Field, Alexander J. 2008. The impact of the second World War on us productivity growth. *Economic History Review* 61(3): 672–694.

Field, Alexander J. 2011. *A great leap forward: 1930s depression and U.S. economic growth*. New Haven: Yale University Press.

Fishback, Price V., and Joseph A. Cullen. 2013. Second World War spending and local economic activity in US counties, 1939–58. *Economic History Review* 66(4): 975–992.

Fishback, Price V., William C. Horrace, and Shawn Kantor. 2005. Did new deal grant programs stimulate local economies? A study of federal grants and retail sales during the great depression. *Journal of Economic History* 65(1): 36–71.

Glaeser, Edward L. 2008. *Cities, agglomeration, and spatial equilibrium*. Oxford: Oxford University Press.

Glaeser, Edward L., and Kristina Tobio. 2008. The rise of the sunbelt. *Southern Economic Journal* 74(3): 610–643.

Goldin, Claudia D. 1980. War in American economic history. In *The encyclopedia of American economic history*, vol. 3, ed. Glenn Porter, 73–140. New York: Scribner's and Sons.

Gordon, Robert J., and Robert Krenn. 2010. The end of the great depression 1939–41: Policy contributions and fiscal multipliers. Technical report, National Bureau of Economic Research.

Haines, Michael R. 2010. Historical, demographic, economic, and social data: The United States, 1790–2000, volume ICPSR02896-v3. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor.

Higgs, Robert. 1987. *Crisis and leviathan: Critical episodes in the growth of American government*. Oxford: Oxford University Press.

Higgs, Robert. 2006. Private profit, public risk: Institutional antecedents of the modern military procurement system in the rearmament program of 1940–41. In *Depression, war, and Cold War: Studies in political economy*, ed. Robert Higgs. Oxford: Oxford University Press.

Jaworski, Taylor. 2015. World War II and the industrialization of the American South, 1930–1960.

Koistinen, Paul A.C. 2004. *Arsenal of World War II: The political economy of American warfare, 1940–1945*. Lawrence: University Kansas Press.

Krugman, Paul. 2011. Oh! What a lovely war! *New York Times*, 15 Aug 2011.

Lewis, Robert. 2007. World War II manufacturing and the postwar southern economy. *Journal of Southern History* 73(4): 837–866.

Markusen, Ann, Peter Hall, Scott Campbell, and Sabina Deitrick. 1991. *The rise of the Gunbelt: The military remapping of industrial America*. Oxford: Oxford University Press.

Nash, Gerald D. 1990. *World War II and the west: Reshaping the economy*. Lincoln: University of Nebraska Press.

Rhode, Paul W. 2003. After the war boom: Reconversion on the pacific coast, 1943–1949. In *History matters: Essays on economic growth, technology, and demographic change*, ed. Timothy Guinnane, William Sundstrom, and Warren Whatley. Stanford: Stanford University Press.

Rockoff, Hugh. 2012. *America's economic way of war: War and the US economy from the Spanish-American war to the Persian gulf war*. Cambridge: Cambridge University Press.

Romer, Christina D. 2011. The hope that flows from history. *New York Times*, 13 Aug 2011.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated public use microdata series: Version 5.0 [machine-readable database]*. Minneapolis: University of Minnesota.

Siu, Henry E. 2008. The fiscal role of conscription in the us world war ii effort. *Journal of Monetary Economics* 55(6): 1094–1112.

Smith, R. Elberton. 1959. *The army and economic mobilization*. Washington, DC: GPO.

U.S. Department of Commerce Bureau of Economic Analysis. 2015. Regional economic accounts.

US Selective Service System. 1948. Special Monograph No. 12. Government Printing Office, Washington, DC.

Vernon, J.R. 1994. World War II fiscal policies and the end of the great depression. *Journal of Economic History* 54(4): 850–868.

White, Gerald T. 1980. *Billions for defense: Government finance by the defense plant corporation during World War II*. Tuscaloosa: University Alabama Press.

The Mirage of the German Armament Miracle in World War II

Jonas Scherner and Jochen Streb

Abstract Based on a revision of output and labor productivity series of the German armament industry during the World War II we show that the so-called German armament miracle was only a statistical construct produced by the manipulation of Albert Speer and his statisticians. In particular, we find no structural break in the year 1942 when Speer became armament minister. In a period of continuous expansion, only the year 1941 stands out, when both domestic armament production and labor productivity decreased. We suggest that this temporary slump was caused by the short-term effects of outsourcing activities and capital deepening.

Keywords Albert Speer • Armament miracle • Germany • Labor productivity • Occupied countries • Outsourcing • Second World War

Introduction

Nine years after the end of the Second World War, the West Berlin-based German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung) published a book about the performance of German industry during the war (Wagenführ 1954). The book was written by Rolf Wagenführ, a statistician who had headed the statistical department of Speer's armaments ministry between 1943 and the end of the war and who continued his career as a general director of the statistical office of the EC and a professor for statistics in Heidelberg (Tooze 2001). A first draft of Wagenführ's book was finished during the last months of the war and the Allies relied on it in their early studies on the performance of the German war economy. The published book, too, became very influential during the following decades. It shaped common understandings of the German war economy and continues to do

J. Scherner
Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
e-mail: Jonas.scherner@ntnu.no

J. Streb (✉)
Universität Mannheim, 68131 Mannheim, Germany
e-mail: streb@uni-mannheim.de

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_10

Table 1 The supposed German armament miracle (New Year 1941–1942 = 100)

	Armament production	Workers	Labor productivity
New Year 1941–1942	100	100	100
New Year 1942–1943	177	113	157
New Year 1943–1944	225	119	189
June/July 44	300	130	234

Source: Wagenführ 1954, p. 125

so until today (USSBS 1976; Eichholtz 1985; Overy 1994; Abelshauser 1998). A second edition was even published in 2006.

One of the most influential pieces of information presented in Wagenführ's book is an index on German armament production during the war. This index was originally prepared on behalf of Albert Speer's armament department and captures armaments in the strict sense such as weapons or combat aircraft but omits consumer goods or uniforms supplied to the Wehrmacht. Measurement is based on the physical volume of those newly produced armament goods that had been delivered to the military procurement agencies. These armaments were assigned to different subgroups (such as ammunition or aircraft) and weighted with the prices of 1943.[1] According to this index German armament production stagnated in 1940 and 1941 but more than tripled between early 1942 and July 1944. What is even more impressive is that this increase was apparently accompanied by a huge increase in labor productivity in the armament industries (Table 1).

It is important to note that the armament department constructed the armament index first and foremost to convince Hitler of Speer's managerial capabilities. In his postwar publication, however, Wagenführ (1954) presents the armament index as hard facts. In his memoirs, Speer (1969) in turn referred back to Wagenführ's statistical results to document his own competence. In spite of the dubious nature of this circular reasoning Speer's alleged managerial capabilities are also at the core of the two most important master narratives of the German armament miracle. The first one of these was the Blitzkrieg-Hypothesis. An early version of this hypothesis was proposed immediately after the war by the *United States Strategic Bombing Survey* (USSBS) and subsequently elaborated by former contributors to this survey, such as Nicholas Kaldor (1946) and Harvard economist Burton H. Klein (1959). Later on, the Blitzkrieg-Hypothesis was refined by the British economic historian Alan Milward (1965). According to the proponents of the Blitzkrieg-Hypothesis, Hitler deliberately decided to under-mobilize the German armament industry during the first half of the war in order to free up resources for producing consumer goods that were needed to maintain the German public approval of the National Socialists' antagonistic policies. Only after the military failures in late 1941 did the German

[1]For more details, see BArch R 3/1732, Planungsamt, Hauptabteilung Planstatistik, Bearbeitung: Maschinelles Berichtswesen, Indexziffern der deutschen Rüstungsendfertigung, Stand Januar 1945, fol. 1–2; Zusammensetzung des Index der Rüstungsendfertigung, fol. 54–59.

military planners acknowledge that they had to considerably increase weapons production by allocating as many resources and workers to armament production as possible. This was made possible only because of Speer's outstanding managerial capabilities. Nicholas Kaldor (1946, p. 48), for example, stated:

> Speer's administration in the course of the following two-and-a-half years was the single great success which the German war economy can record, and the only that will retain a more than historical interest.

Since the 1980s, however, the validity of the Blitzkrieg-Hypothesis has been challenged by proponents of a rival hypothesis, the Inefficiency-Hypothesis. According to Richard Overy (1994) it was the untimely outbreak of the Second World War, originally planned to start no earlier than 1942, which put German mobilization activities in temporary disorder. In a variation stressed by Rolf-Dieter Müller (1988), the armament industry's inefficiency during the first half of the war mostly resulted from the political incoherence (polycracy) of the "Third Reich". Overy and Müller agree that it was Speer who solved many organizational problems of the early German war economy. Proponents of both the Blitzkrieg-Hypothesis and the Inefficiency-Hypothesis credit Speer with having made the following political decisions (USSBS 1976, p. 7; Abelshauser 1998; Overy 1994, pp. 356–363; Weyres-v. Levetzow 1975, pp. 47–49). First, the number of weapon types was reduced, which might have allowed many companies to move to mass production and exploit economies of scale. Second, the frequency of minor design changes was decreased, so firms could save some of the costs that came along with the adaption of their production processes. Third, against the declared desire of the armed forces, finishing procedures like polishing or lacquering that add nothing to the destructive power of a weapon were abolished, which reduced the working hours needed to produce one piece of an armament good. Fourth, firms were forced to share technological know-how in newly established inter-firm committees in order to give less-efficient firms the information considered necessary for imitating the technology of the superior firms. This policy might have especially accelerated the diffusion of flow production techniques in German industry. Fifth, it is assumed that under Speer's reign an increasing share of military orders were placed to the most-efficient armament manufacturers. Most importantly, Speer is praised for having replaced the traditional regime of cost-plus contracts with fixed-price contracts that in contrast to the former motivated companies to increase their efficiency (Overy 1994, p. 357).

Yet, recent studies suggest that the very positive evaluation of Speer's managerial capabilities might be misleading. First of all, the number of weapon types was actually not reduced before 1944 in the most important armament sub-groups, such as, for example, aircraft (Scherner and Streb 2006). With respect to knowledge exchange, business studies imply that also this measure was rather unsuccessful (Scherner and Streb 2008). In addition, it was very difficult for the German procurement agencies to identify and utilize the most efficient producers, at least in those cases when firms manufactured complex armaments with the help of a large network of specialized suppliers (Scherner 2015). Other rationalization measures attributed to Speer had been implemented long before his appointment. This applies most prominently to the incentive structure. Already during the late 1930s fixed-price contracts had been in common use (Rauh-Kühne 2002; Scherner 2008; Budraß et al. 2010). It is thus not surprising that labor productivity on the companies' level had begun to increase already before Speer's appointment, as recent microeconomic studies show (Scherner and Streb 2010). In addition, as Adam Tooze (2005) emphasizes, Speer can hardly be credited with the rapid increase of the armament production by 50 % in the first half of 1942 due to the considerable time lag between the start of production of an armament good and its delivery to the procurement agencies of the Wehrmacht. In the case of aircraft, for example, this production span amounted to up to 9 months.

Whereas all these arguments raise serious doubts about whether Speer's alleged reforms were actually responsible for the increase in armament output and labor productivity beginning in 1942, this article focuses on several deficiencies of the macroeconomic data underpinning the so-called armament miracle. One major problem of the published armament index is that Wagenführ provides no monthly data on armaments and labor productivity for the period before January 1942. In other words, we have no clear idea of what exactly was going on in the German armament industry before Speer became minister. A second major problem stems from the fact that the armament index also includes armament goods that were produced outside of Germany, especially in occupied countries. In contrast, the workforce numbers Wagenführ uses to calculate labor productivity capture only the workers in the armament industries within the German borders (Scherner and Streb 2006, p. 178; Wagenführ 1954, p. 125). As a result, labor productivity growth in armament production is systematically overestimated. Thus, the main agenda of this chapter is to present a revision of the armament and labor productivity indices. In a first step, we add information about the period September 1939 to December 1941 to Wagenführ's index, and provide some crosschecks. Secondly, we subtract armament goods that were supplied by occupied countries and provide a real "German" armament index. Relying on this, we then revise the labor productivity series. Finally, we will try to explain the observed U-shaped development of labor productivity.

A Reconstruction of the Armament Index for the Period September 1939 to January 1942

Wagenführ provides in his book monthly data for the armament index as well for the sub-group indices only from January 1942 onwards. Yet, it is possible to reconstruct these data for the early war period. In a first step we collected scattered data about the sub-group indices for the first half of the war, which are shown in Table 2. For ammunition, tanks, and war vessels, information about monthly output between September 1939 and the end of 1941 can be found in the files of the armament

Table 2 Sub-group indices September 1939 to December 1941, Jan 42 = 100

Month	Weapons	Tanks	War vessels	Powder and explosives	Aircraft	Ammunition
Sep 1939	63	7	55	51	38	90
Dec 1939	63	20	32	51	38	82
Mar 1940	68	27	17	59	64	86
Jun 1940	79	33	24	97	97	149
Sep 1940	83	44	33	107	119	112
Dec 1940	86	60	76	94	71	98
Mar 1941	114	53	64	104	119	112
Jun 1941	123	73	92	113	104	107
Sep 1941	88	86	118	109	104	91
Dec 1941	83	100	138	112	104	100

Source: See text

ministry, in Wagenführ's private files,[2] and in a report of the USSBS.[3] The same sources offer data on aircraft production between January 1940 and December 1941. For weapons, ammunition, powder, and explosives, the USSBS (1976, p. 283) provides only quarterly data covering the period between October 1939 and the end of 1941. To reconstruct the monthly figures we assume that the monthly index equals the respective quarter value. For these sub-groups, we also assign the index value of the last quarter of 1939 to the monthly index of September 1939. In the case of the aircraft production, the value of January 1940 has to be used to estimate the monthly index for the last 4 months of 1939. In the case of truck production, known to be of very small relative size, for which no data could be discovered, we assume a constant monthly production amounting to the index level of 100 (USSBS 1976, p. 144).

In order to overcome the problem that Wagenführ reveals no information about the weights of the sub-group indices of his armament index, we first convert the monthly group index numbers to real value variables by using information on the production value of the sub-groups in January 1942 in constant prices of 1943 that are available in his book. In a second step, these monetary values were added up for each month. Finally, the resulting monthly time series of the total real value of armament production from September 1939 to January 1942 is transformed into an index based on January 1942. From a theoretical perspective, this method leads to the same result as if we had used the unknown weights.

[2]These files are part of the empirical base of Wagenführ's book and are now held at the *Wirtschaftsarchiv Baden Württemberg* (WIBAWU). Wagenführ, who was a professor at the University of Heidelberg after the Second World War, gave these files to the economic historian Eckart Schremmer who in turn left them with Jochen Streb in 2003.

[3]BARch R 3/3005 (tanks, war vessels, aircraft production); WIBAWU N10/29, HA Planstatistik Index der industriellen Rüstungsproduktion, Geheime Reichssache, 6 Ausfertigungen, (aircraft production).

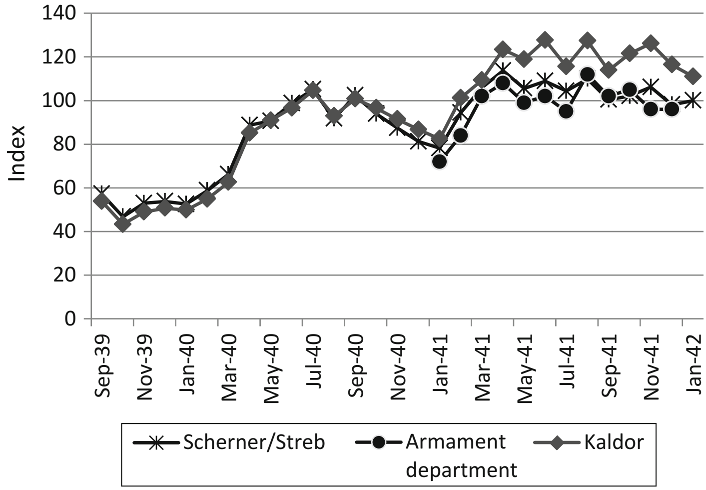


Fig. 1 Completing and cross-checking Wagenführ's armament index, September 1939–January 1942 (Source: see text)

Table 3 Kaldor's weighting scheme

Sub-group	Weighting factor
Tanks	11
Weapons	9
Motor vehicles	3
Aircraft	42
Naval construction	7
Ammunition	25
Powder	1
Explosives	2

Source: BIWAWU N10/29

Figure 1 shows the armament index, marked with asterisks, we reconstructed by using the method and data described above. Two cross-checks suggest that our estimation of the armament index is reliable. The first cross-check is based on sub-groups' weighting factors suggested by Nicholas Kaldor, which we discovered in Wagenführ's inheritance and which are presented in Table 3. The line marked with squares shows the development of the armament index that results if we aggregate the group indices with Kaldor's weighting scheme. In the year 1940, this index corresponds to the index resulting from our method, but in 1941 it leads to a higher level of total armament production. The second cross-check rests on monthly figures of the total armament index for 1941 (in million Reichsmark (RM), prices of 1943) Burton H. Klein provides in a publication on the German war economy. Klein's figures (1959, p. 190) are identical with monthly data for the same year that we discovered in the files of the armament ministry (BARch R 3/3005, fol. 68). These

Table 4 Annual armament production, billion Reichsmark, in prices of 1943

	Wagenführ series	Annual increase	Scherner/Streb series	Annual increase
1939	6.25		6.25	
1940	11	76 %	10	60 %
1941	11	0 %	12	20 %
1942	16	45 %	16	33 %

Source: See text

data are also shown in Fig. 1, marked with dots, and confirm our estimate. Note that our findings are also in line with the estimates Tooze (2006, p. 348) presents only graphically.

Table 4 shows that Wagenführ's average production figures suggest that armament production stagnated in 1941 and sharply increased in 1942 by 45 % (Wagenführ 1954, p. 125) whereas our estimate implies a more continuous development with growth rates of 20 % and 33 %, respectively. We can only speculate why Wagenführ presented the wrong figures: either he made a massive calculation error or he deliberately provided incorrect values. The latter appears more likely because a stagnation of armament production in the first half of the war fits neatly into the general narrative of success which he constructed for Speer. What other reasons might Wagenführ have had not to provide monthly index numbers for the period before January 1942, although he possessed the respective data for at least some group indices, as both his personal papers and a contemporary USSBS report reveal? And why were the monthly data about total armament production for 1941 not included in a report about the German war economy sent by the armament ministry to Nazi leaders in July 1944? In particular, why did this report only cover the period when Speer headed the armament ministry in spite of the fact that all other sectors of the economy were covered in full (Scherner 2007)?

Taking into Account Foreign Armament Production for Germany

As mentioned above, Wagenführ's armament index includes the foreign production of armaments delivered to the Wehrmacht, which was especially carried out in occupied countries. Yet, until recently, no quantitative information about the extent of foreign deliveries was available. Scherner (2012b) overcomes this shortcoming by estimating the monthly foreign armament production during the war period. Foreign production for the Wehrmacht can be subdivided into armament goods in a narrow sense such as components needed to produce ammunition, tanks, or aircraft that were included in the armament index, and other goods such as uniforms, steel helmets, or office furniture. Relying on fragmented evidence Scherner (2012b) assumed that armament goods in a narrow sense had a constant share of about two-thirds in total foreign production delivered to the Wehrmacht. Considering that the Reichsmark was overvalued in most occupied countries, he first converted the data on foreign armament production to nominal RM-values by using purchasing power parities calculated by the Research Center on the Military Economy (Forschungsstelle für Wehrwirtschaft). In a second step, these nominal values were transformed in real values at 1943 RM-prices by relying on the price index for armament goods Karl Hettlage provided in summer 1945 to the USSBS.[4] According to this method, between 1940 and 1944, the most important foreign producers of armament goods in a narrow sense were the Protektorat (Czech rump state) (8.7 billion RM), France (6.2 billion RM), Belgium (3.2 billion RM), Netherlands (2.2 billion RM), and the Generalgouvernement (parts of Poland) (1.4 billion RM).

By subtracting the respective foreign armament production from the value of total armament production captured by the armament index, we are in principle able to calculate an armament index for the military production within the German borders of 1939. One problem, however, remains. Occupied countries mostly produced components, parts, and intermediate products whereas the armament index only covers end products. This implies that we cannot assume that the supply of foreign products had an immediate effect on the final output of German armaments in the same month in which they were delivered. Unfortunately, we do not have exact information on the time spans different armament industries needed to assemble the foreign deliveries to final products. That is why, for the period from 1941 to 1943, we allocate annual foreign production evenly to the monthly German armament production of the respective calendar year. For the years 1940 and 1944, we take into account that in these years most of the foreign deliveries stemmed from the second (1940) or the first (1944) half of the calendar year. The fourth row of Table 5 specifies this allocation scheme.

We observe both the largest increase in foreign production and the peak of its share in total armaments output in the year 1941. In this year, more than one quarter of German armament output was manufactured in foreign territories. This finding challenges the conventional wisdom which suggests that occupied countries were fully and systematically exploited only after Speer's appointment as armament minister (Milward 1966, p. 79, 109). This result rather fits into a recent study which reveals that the incentive structure and the regulatory framework for the placement of orders to occupied countries was mainly created before Speer became armament minister (Dickert 2014). It also supports a recent re-estimate of total

[4]Karl Hettlage was the former head of a division of the German armament ministry who had been responsible for price checks of armament producers. According to Hettlage's price index, armament prices increased both in 1941 and 1942 and dropped significantly later. National Archives and Record Administration (NARA) RG 243, United States Strategic Bombing Survey, Overall Economic Effects Division, Special Paper No 1, The Gross National Product of Germany 1936-1944, pp. 5–9. On Hettlage's tasks in the armament ministry, see BArch R 8136/3732, Neugliederung des RMRuK, 22.9.1943. Similar to Wagenführ, Hettlage continued his career after World War II as a law professor, as a state secretary in the Federal Ministry of Finance (Bundesfinanzministerium), and as president of the Munich-based Ifo-Institute. About Hettlage, see Schrafstetter (2008).

Table 5 Foreign armament production, million RM, in 1943 prices (PPP)

	Foreign armament production	Share in total armament production (%)	Allocation scheme
1940	599	6	Jan.–Jun: 7 million RM
			Jul. –Dec.: 93 million RM
1941	3328	28	Jan.–Dec.: 277 million RM
1942	3958	24	Jan.–Dec.: 330 million RM
1943	4425	17	Jan.–Dec.: 369 million RM
1944	5454	17	Jan.–Jun.: 606 million RM
			Jul. –Dec.: 303 million RM

Sources: See text and Scherner (2012b, Table 6)

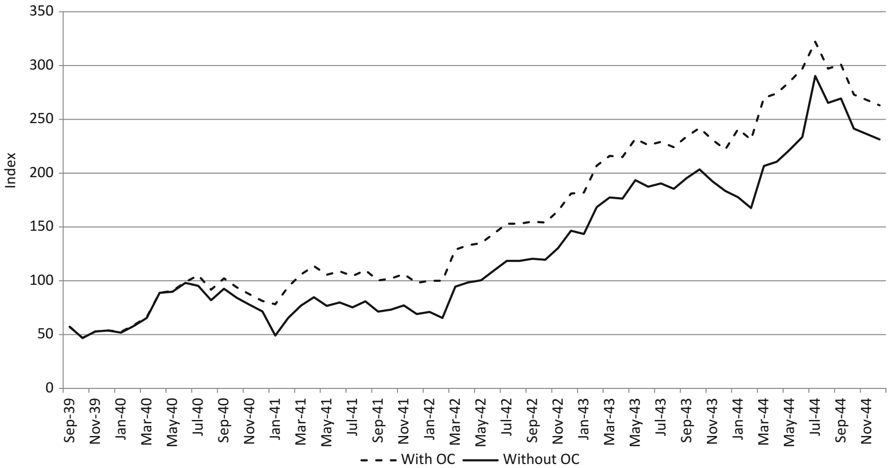


Fig. 2 German armament production with and without the deliveries of occupied countries (Sources: See text. *Notes*: German plus foreign production Jan 42–Feb 42 = 100)

German real imports showing that the largest increase can be observed between 1940 and 1941, followed by a slower growth in the subsequent years (Scherner 2012a). All these observations lead to the conclusion that Speer did not revolutionize foreign procurement but mostly relied on existing practices already established in the first half of the war.

Based on the allocation scheme as suggested in Table 5, we construct an armament index that covers only those weapons that were actually produced within the borders of Germany of 1939. Figure 2 reveals the striking result that it took almost 2 years before German armament production (without the contribution of foreign countries) reached the same level as already in the summer of 1940. Still more surprising is the fact that at its lowest level, in January 1941, the German armament industry did not produce more than about 50 % of the peak in 1940.

The fact that it was only 1941 that was different is also confirmed by the annual development of German war production without foreign deliveries, which is

Table 6 Annual German armament production, billion Reichsmark, in prices of 1943

	Wagenführ series	Annual increase	Scherner/ Streb-series with OC	Annual increase	Scherner/ Streb-series without OC	Annual increase
	6.25		6.25		6.25	
1940	11	76 %	10	60 %	9.13	46 %
1941	11	0 %	12	20 %	8.41	−8 %
1942	16	45 %	16	33 %	12.36	47 %

Source: see text

presented in row 6 of Table 6. Whereas the German armament production increased both in 1940 and in 1942 by almost 50 %, it even decreased in 1941 by 8 %. The reasons for this slump will be explained in the next section.

Calculating Labor Productivity of German Armament Production

When writing his famous book Wagenführ was probably well aware that his calculation of labor productivity (armament production/armament workers), covering the period from the turn of the year 1941–1942 up to 1944, was heavily distorted. First, the numerator was too large because of the inclusion of production in the occupied territories. Secondly, the denominator was too small because it did not account for most of the workforce in the supplier industries that contributed significantly to total armament production. Wagenführ explicitly referred in his book to the numerator problem. As his private files make clear, he was also aware of the denominator problem.[5]

In Fig. 3, we consider only the production in the prewar territory of Germany, excluding the Protectorate and the General Government. We use, like Wagenführ, the workforce figures of the so-called A-firms (armament firms), excluding, however, again the workforce of A-firms in the Protectorate and General Government, which are included in his figures. Note that we also overestimate labor productivity growth after 1941 because our denominator, the labor input, is still too small because of two reasons. First, despite the growing importance of armament firms'

[5]WIBAWU N10/29, Dr. Fritz Grotius, 8.10.1946, Betrifft: Arbeitsproduktivität. Statistik über Arbeitseinsatz, Bezug: Fragen von Dr. E. C. R. Reeve, B.B.S.U, Air Ministry, London. In this letter, sent from Dr. Fritz Grotius, a former colleague of Wagenführ in the planning office of the armament ministry, to Dr. Reeve of the British Air Ministry in 1946, Grotius emphasized that the workforce data underestimated the size of the "true" number of workers employed in the armament industry due to the omission of significant parts of the supplier industry.

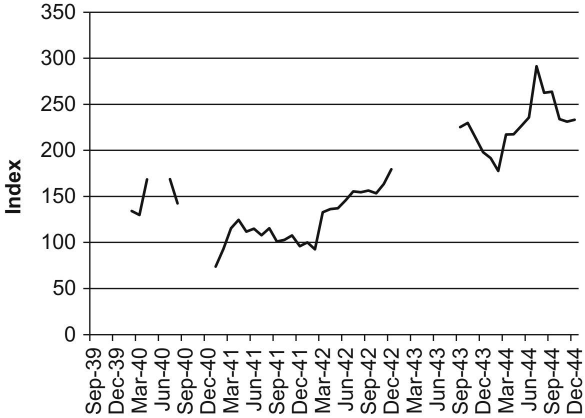


Fig. 3 Productivity miracle? A matter of perspective (Sources and notes: For monthly employment data in 1940, see Militärgeschichtliches Forschungsamt (2000, p. 893, 902). For monthly employment data in 1941 and 1942, see BArch R 3/1804 and BArch R 3/1805. Note that we subtracted the employment figures of the Rüstungsinspektion 38, i.e., the Protectorate, and of the Rüstungsinspektion 39, i.e., the Generalgouvernment, in order to capture only the armament workers within the German borders. For monthly employment data for June 1943 as well as for the period between September 1943 and December 1944, see USSBS (1976) Appendix Table 10, p. 212. A cross-check with data provided in BARch R 3/1965, fol. 115 and fol. 119, suggests that the USSBS data for this last period of the war do not include armament workers both of the Rüstungsinspektion 38 and the Rüstungsinspektion 39)

outsourcing activities, only a smaller part of their suppliers is listed in the A-firms sample. As a result we miss many workers from non-A-firms who produced parts and components of armament goods. Second, the average weekly working hours increased significantly compared to the previous years (Boelcke 1969, p. 12). This is especially true for 1944.

Figure 3 strongly questions the conventional wisdom that something like an armament miracle ever existed, because the “productivity miracle” of the German war economy (see Table 1) largely disappears once we use the adjusted data. In December 1942, labor productivity is only slightly higher than at its early peak value in summer 1940. Speer’s “productivity miracle” of 1942 only occurs because of Wagenführ’s decision to select January and February 1942 as the basic months of comparison. In these months, the productivity level was exceptionally low – also due to seasonal reasons, because during the winter aircraft could not be flown in and therefore not be accepted by the procurement agencies (Scherner and Streb 2006, p. 177). Taking seasonal factors into account, we see that labor productivity in December 1943 was only 10 % higher than 1 year earlier. Even the maximum peak in labor productivity which was reached in June of 1944 exceeds the mid-1940 levels only by 70 %. The question remains, however, why labor productivity dropped in 1941.

Explaining the U-Shaped Labor Productivity Growth

We assume that industry-specific learning curves played a major role in explaining why labor productivity in German armaments production was comparatively low between summer 1940 and March 1942. The idea of learning curves is based on the empirical observation that the direct amount of labor required for manufacturing one unit of an armament good regularly declines when the total output of this type is expanded. This relationship can be graphically expressed by the so-called learning curve (Budraß et al. 2010, p. 129). The basic explanation for the negative slope of this function is that workers learn as they work. In this respect, learning-by-doing means that the more often a worker repeats a special task, the more efficient he or she will become. An alternative explanation is that the learning effects are embodied in the production system, which is controlled by the managers, and not in the individual workers. In this respect, we can think of technical rationalization measures, the refining of the production methods or the introduction of assembly lines. Learning effects might arise in all kinds of industries, but the increase in labor productivity is expected to be especially high when workers are given rather complex tasks.

These considerations have two major implications. We expect *average* labor productivity growth to slow down if at least one of the following conditions holds. First, the focus of armament production shifts from industries with fast learning effects to industries with slow learning effects. Second, many firms without previous experience in the production of armament goods turn towards the manufacturing of weapons. Both occurred in 1940–1941.

Up to the summer of 1940, the expansion of armament production was primarily driven by an increase in ammunition production (see Table 2) which is a sub-group of armament goods in which learning effects could be relatively quickly realized (Scherner and Streb 2010). Thereafter, it was the aircraft industry that massively expanded. This industry produced technologically complex products and therefore needed much more time to move down its learning curve.

Generally, we can observe a massive increase in both the number of companies producing armament goods and the armament industry's capital stock after the summer of 1940, as shown in Table 7. The investment in newly founded armament

Table 7 Annual growth rates of the number of German A-firms and the capital stock of the German armament industry, in percent, 1939–1944

	1939	1940	1941	1942	1943	1944
Capital stock	48	71	42	31	29	n.a.
A-firms	44	69	17	8	1	26

Sources: Growth of capital stock is calculated from investment figures in armament plants provided by Scherner (2010). It is assumed that before 1934, given the restrictions of the Versailles treaty, no significant production capacities for armaments existed, and that the average annual depreciation rate amounted to 10 %. Data about A-firms are taken from Scherner et al. 2014, p. 1001 firms combined with the hiring of a large, new, and inexperienced workforce certainly decreased overall labor productivity in 1941. The long-term effect was the opposite: As the ratio between experienced and inexperienced companies, workers, and managers increased after 1941, labor productivity growth again accelerated because of learning effects. Hence, it was not necessarily Speer's remaining reforms (such as administrative rationalization) that explain the productivity growth of the second half of the war but rather the fact that the many armament firms that had been established at the beginning of the war almost naturally moved down their learning curves from 1942 onwards.

Note, however, that the number of A-firms shown in Table 7 represents only a lower limit of all German companies involved in armament production because of the increasing supplier network during the war (Scherner et al. 2014). The political decision to support the deepening of the inter-firm division of labor can be understood as an answer to the failure of the originally preferred centralization strategy. Before the war, armament production had been partly concentrated in central Germany because this area seemed to be out of the range of potential enemy aircraft given their operating distance at the time. Yet, after the war had started, the growing military demand for armament goods produced by the firms in central Germany soon exceeded their production capacities. The Nazi economic planners first thought that the resulting labor shortage at the central German production sites could be resolved by making use of workers of metalworking companies located in the western border regions. Based on various regulations enacted in 1938 and 1939, authorities commanded these workers to move and work at the armament firms in central Germany. But armament manufacturers were not at all satisfied with the performance of the "forced" German workers. Moreover, regional Nazi leaders of border regions complained about the lack of military orders and the rationing of raw materials for civilian production which threatened to shut down metalworking companies in these areas. These complaints led politicians to change to the new outsourcing strategy: At the end of 1939, it was decided to spread future procurement orders all over Germany in order to keep the medium-sized firms in business during wartime. What is more, large German armament firms were asked to use small and medium-sized firms as sub-contractors.

Armament firms had good reasons to obey this wish of the regime willingly. Facing the excess demand for armament goods, firms had to choose between one of the three following alternatives. First, they could try to decline some of the actual orders. This strategy, however, came not only with the risk that military procurement agencies would disregard them when allocating future orders across armament firms but also meant that the firms had to forgo the usually high profits of state-financed armament production. Second, armament firms could decide to build up additional production capacities by investing in new plants and machinery. This alternative had its shortcomings too. Often enough, firms were not willing to take the risk of investing in what they assumed would become excess capacity once the war ended and, consequently, of little long-term value. The solution to these problems was to rely on the new third alternative: By outsourcing the production of intermediate and final goods, armament producers could use underutilized capacities of firms from

Table 8 Average annual growth rates of outsourcing ratios in the Ju 88 program

Production stage	Annual growth rate, 1937–1943 (%)
Final assembly	13
Aircraft engines	21
Components of aircraft engines	2
Components of airframes	19
Light metal	4

Outsourcing ratio: intermediate goods/production costs
Source: Scherner et al. (2014, Table 7)

other industries, and at the same time avoid investments in capital goods which could be fully utilized only under the conditions of war.

Outsourcing played an important role especially in the production of complex armaments, as shown in the case of the Ju 88-program (Budraß et al. 2010; Scherner et al. 2014) which was established in 1938 and which constituted one of the largest German armament projects. The firms that were involved in this program employed between one-third and about one-half of the German aircraft production workforce during the following 6 years. During this period, some 14,000 Ju 88 bombers were built. A detailed analysis of the outsourcing activities of the firms involved in this program shows staggering outsourcing rates, as presented in Table 8.

The outsourcing activities had again opposite short- and long-term effects. In the short term, they lowered average labor productivity because the newly added suppliers of parts and components had to solve many initial production problems, such as the implementation of the know-how transfer, training the workers, and the changing of production installations. In the long term, however, the increased inter-firm division of labor led to "economies of scale" and "economies of learning" and thus to the increase in labor productivity observable from early 1942 onwards.

Conclusions

The mirage of the so-called German armament miracle was evoked by the manipulations and omissions of armament minister Albert Speer and his statistician Rolf Wagenführ: first of all by Wagenführ's manipulation of the production figures of 1940 and 1941, which makes Speer's early achievements appear much more impressive, and, secondly, by the decision not to publish existing information on the early war years, which obscures this period to this day. After taking into account the contribution of the occupied countries, our revision of the armament index reveals a rather continuous growth of armament output and labor productivity during the Second World War. Especially, we find no structural break in the year 1942 when Speer became armament minister. Only the year 1941 stands out negatively when both the sudden shift to aircraft production and the integration of many unexperienced armament manufacturers led to a considerable decrease in labor productivity. Hitler started his invasion of Russia exactly at the moment when his armament industries were confronted with huge temporary transformation problems caused by the short-term effects of outsourcing activities and capital deepening.

References

Abelshauser, W. 1998. Germany: Guns, butter and economic miracles. In *The economics of World War II: Six great powers in international comparison*, ed. M. Harrison, 122–176. Cambridge: University Press.

Boelcke, W.A. 1969. *Deutschlands Rüstung im Zweiten Weltkrieg. Hitlers Konferenzen mit Speer*. Frankfurt: Athenaion.

Budraß, L., J. Scherner, and J. Streb. 2010. Fixed-price contracts, learning and outsourcing: Explaining the continuous growth of output and labour productivity in the German aircraft industry during World War II. *The Economic History Review* 63: 107–136.

Dickert, E. 2014. Die "Nutzbarmachung"des Produktionspotentials besetzter Gebiet durch Auftragsverlagerung im Zweiten Weltkrieg. Staatliche Regulierung und Verlagerungsverhalten von Maschinenbau- und Automobilunternehmen. Dissertation, Trondheim.

Eichholtz, D. 1985. *Geschichte der deutschen Kriegswirtschaft*, vol. 2. Berlin: Akademie Verlag.

Kaldor, N. 1946. The German war economy. *The Review of Economics and Statistics* 13: 33–52.

Klein, B.H. 1959. *Germany's economic preparations for war*. Cambridge MA: Harvard University Press.

Militärgeschichtliches Forschungsamt. 2000. Germany and the Second World War, Vol. 5: Organization and Mobilization of the German Sphere of Power, Part I: Wartime Administration, Economy, and Manpower Resources, 1939–1941, Oxford: University Press.

Milward, A.S. 1965. *The German economy at war*. London: Athlone Press.

Milward, A.S. 1966. *Die deutsche Kriegswirtschaft 1939–1945*. Stuttgart: Deutsche Verlags-Anstalt.

Müller, R.D. 1988. Die Mobilisierung der Deutschen Wirtschaft für Hitlers Kriegsführung. In *Das Deutsche Reich und der Zweite Weltkrieg*, vol. 5/1, ed. Militärgeschichtliches Forschungsamt, 349–692. Stuttgart: Deutsche Verlags-Anstalt.

Overy, R.J. 1994. *War and economy in the Third Reich*. Oxford: University Press.

Rauh-Kühne, C. 2002. Hitlers Hehler? Unternehmerprofite und Zwangsarbeiterlöhne. *Historische Zeitschrift* 275: 1–55.

Scherner, J. 2007. Bericht zur deutschen Wirtschaftslage 1943/44. Eine Bilanz des Reichsministeriums für Rüstung und Kriegsproduktion über die Entwicklung der deutschen Kriegswirtschaft bis Sommer 1944 (Dokumentation). *Vierteljahr Zeitges* 55: 499–546.

Scherner, J. 2008. *Die Logik der Industriepolitik im Dritten Reich. Die Investitionen in die Autarkie- und Rüstungsindustrie und ihre staatliche Förderung*. Stuttgart: Steiner Verlag.

Scherner, J. 2010. Nazi Germany's preparation for war: Evidence from revised industrial investment series. *European Review of Economic History* 14: 433–468.

Scherner, J. 2012a. Der deutsche Importboom während des Zweiten Weltkriegs. Neue Ergebnisse zur Struktur der Ausbeutung des besetzten Europas auf der Grundlage einer Neuschätzung der deutschen Handelsbilanz. *Historische Zeitschrift* 294: 79–113.

Scherner, J. 2012b. Europas Beitrag zu Hitlers Krieg. Die Verlagerung von Industrieaufträgen der Wehrmacht in die besetzten Gebiete und ihre Bedeutung für die deutsche Rüstung im Zweiten Weltkrieg. In *Europäische Volkswirtschaften unter deutscher Hegemonie*, ed. C. Buchheim and M. Boldorf, 70–92. Munich: Oldenbourg Verlag.

Scherner, J. 2015. Die Grenzen der Informationsbeschaffung und -verarbeitung in der deutschen Wehr- und Kriegswirtschaftsverwaltung im Dritten Reich. *Jahrbuch für Wirtschaftsgeschichte* 56: 99–135.

Scherner, J., and J. Streb. 2006. Das Ende eines Mythos? Albert Speer und das so genannte Rüstungswunder. *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte* 93: 172–196.

Scherner, J., and J. Streb. 2008. Wissenstransfer, Lerneffekte oder Kapazitätsausbau? Die Ursachen des Wachstums der Arbeitsproduktivität in den Werken der Deutschen Sprengchemie GmbH, 1937–1943. *Zeitschrift für Unternehmensgeschichte* 53: 100–122.

Scherner, J., and J. Streb. 2010. Ursachen des Rüstungswunders in der Luftrüstungs-, Pulver- und Munitionsindustrie während des Zweiten Weltkriegs. In *Rüstung, Kriegswirtschaft und Zwangsarbeit im Dritten Reich*, ed. A. Heusler, M. Spoerer, and H. Trischler, 37–62. Munich: Oldenbourg Verlag.

Scherner, J., J. Streb, and S. Tilly. 2014. Supplier-networks in the German aircraft industry during World War II and the effects for West Germany's car industry in the 1950s. *Business History* 56: 996–1020.

Schrafstetter, S. 2008. Verfolgung und Wiedergutmachung. Karl M. Hettlage: Mitarbeiter von Albert Speer und Staatssekretär im Bundesfinanzministerium. *Vierteljahr Zeitges* 56: 431–466.

Speer, A. 1969. *Erinnerungen*. Berlin: Propyläen.

Tooze, A. 2001. *Statistics and the German State 1900–1945: The making of modern economic knowledge*. Cambridge: University Press.

Tooze, A. 2005. No room for miracles: German industrial output in World War II reassessed. *Geschichte Gesellschaft* 31: 439–464.

Tooze, A. 2006. The wages of destruction: The making and breaking of the Nazi economy. London.

United States Strategic Bombing Survey (USSBS). 1976. *The effects of strategic bombing on the German war economy overall economic effects division, October 31, 1945*, vol. 1. New York: Garland Publishers.

Wagenführ, R. 1954. *Die deutsche Industrie im Kriege 1939–1945*. Berlin: Duncker & Humblot.

Weyres-v. Levetzow, H.J. 1975. Die deutsche Rüstungswirtschaft von 1942 bis zum Ende des Krieges. Dissertation, Munich.

Neutrality in War

Eric Golson

Abstract Neutrality has long been seen as impartiality in war and is codified as such in The Hague and Geneva Conventions. This chapter investigates the activities of three neutral states in the Second World War and determines, on a purely economic basis, that these countries actually employed realist principles to ensure their survival. Neutrals maintain their independence by offering economic concessions to the belligerents to make up for their relative military weakness. Despite their different starting places, governments, and threats against them, Spain, Sweden, and Switzerland provided similar types of political and economic concessions to the belligerents.

Keywords Economic warfare • Neutrality • Second World War • Realism • Economic concessions

> Of all the neutrals, Switzerland had the greatest right to distinction. She has been the sole international force linking the hideous-sundered nations and ourselves. What does it matter whether she has been unable to give us the commercial advantages we desire or has given too many to the Germans, to keep herself alive? She has been a democratic state, standing for freedom in self-defence among her mountains and in thought, despite of race, largely on our side.[1]

As these words of Winston Churchill suggest, merely through their survival as non-belligerents, neutral countries are honorable. There has been a significant debate about this over the last few decades, with some authors calling the neutrals' actions in both wars despicable and others like Churchill seeing them as rational economic and political actors who sought primarily to preserve their own interests. Using new economic figures and certain Samuelsonian assumptions about causation, this chapter follows Churchill's views, asserting the neutrals were generally rational

[1] Winston Churchill, *The Second World War*: *triumph and tragedy*, vol. VI (1953), p. 712.

E. Golson (✉)
University of Surrey, Guildford, Surrey, UK

School of Oriental and African Studies, University of London, London, UK
e-mail: eg23@soas.ac.uk

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9_11

actors in times of war. Although the specific focus will be on the Second World War because of available data, many of the same trends apply to other wars and these will be mentioned.

The results of this chapter suggest the three European neutrals in the Second World War used *realpolitik* or economic realism to survive the war, adapting their strategies in order to survive. Contrary to the popular literature, from an economic perspective, the neutrals were not pro-German in the face of military pressure. All three actually made a specific effort to avoid being considered as such, using illicit trade and other means to maintain trade links with the Allies. They provided advantages in their terms of trade and capital flows only where necessary to maximize their own utility (in this case, to survive and sustain their populations). By linking the economic outcomes to the changing events in the war, it is possible to confirm neutrality is realism in this period. While certainly not the last word on the subject, the model and new statistics presented herein provide a new and unique view.

World War Two and the Cold War were watersheds for neutrality. Dozens of books appeared in the mid-1990s accusing Spain, Sweden, Switzerland, the Vatican, and others of aiding Hitler's war machine, but few give a particularly accurate view of the political economy and realist choices that countries have to make in war.[2] Survival is essential. It is easy to be impartial if the war is taking place halfway around the world, but it is difficult to ignore if it is happening on your borders. Countries like Finland, under constant threat in the Cold War, struggled to appeal to both sides without gaining much; meanwhile others like Egypt were able to appeal to both sides and gain considerably. Ironically, despite their position it is Finland which reaps the most scorn.[3] Even countries like Jordan in the First Persian Gulf War, whose government was pro-Allied and helped to support the Allies consistently throughout the war, faced considerable and perhaps unwarranted criticism for its neutrality.[4]

The misunderstanding of neutrality became acute in the late 1990s when the United States Senate held hearings into missing assets and the economic gains made by the neutrals in the Second World War, threatening diplomatic retaliation for actions of governments 50 years before.[5] The final Eizenstat report, named for the

[2]The most important texts presenting the revisionist position on Switzerland are Mark Aarons and John Loftus, *Unholy trinity* (1998); Tom Bower, *Blood money* (1997); Adam LeBor, *Hitler's secret bankers* (1997); Werner Rings, *Raubgold aus Deutschland*: *die Golddrehscheibe Schweiz in Zweiten Weltkrieg* (1990); Philipp Sarasin and Regina Wecker, *Raubgold Réduit Flüchtlinge* (1998); Isabel Vincent, *Swiss Banks*, *Nazi gold*, *and the pursuit of justice* (1997); and Jean Ziegler, *The Swiss*, *the gold*, *and the dead* (1998).

[3]Jussi M. Hanhimaki, "Non-aligned to what?" in Bott, Hanhimaki et al. (eds.), *Neutrality and neutralism in the global cold war*: *between or within the blocs*?"(2016)

[4]"Jordan ends neutrality, assailing allied war effort," *The New York Times*, February 7, 1991.

[5]See United States Congress: Senate. Committee on banking, housing, and urban affairs, Swiss banks and the status of assets of Holocaust survivors or heirs: hearing before the Committee on banking, housing, and urban affairs, United States Senate, one hundred fourth congress, second

leader of the United States Commission investigating the neutrals, put considerable emphasis on the movements of gold between the Third Reich and neutral countries as well as Jewish deposits in Swiss banks. Subsequent reports of gold transfers were prepared by several European Central banks.[6] Although the movement of gold in the Second World War is important, there are underlying merchandise trade, services, and payments reasons for these gold payments which are not explored in these reports. The movement of gold is only a small part of a much larger story about neutral–belligerent relations.[7] One of these authors, Professor Jean Ziegler, goes so far as to charge the Swiss government of being more docile than the Vichy French government in acquiescing to German demands.[8] The arguments presented in this chapter do not suggest this is true. The debate has colored the entire discussion of neutrality without looking at the facts.

In order to understand what neutrality is and has traditionally been, it is necessary to review the arguments behind neutrality. Legal scholar Stephen Neff argues impartiality is a modern phenomenon which imposes obligations on those who want to remain at peace.[9] An entire chapter of Neff's work, *The Rights and Duties of Neutrals*, overviews the many rules created during the early twentieth century to protect neutrals.[10] Legal scholars see these treaties as valuable in protecting neutral rights during the war.

But the treaties only provide the framework for neutral–belligerent relations. Treaties did not cover advances in weapons and technology. Neutrals found it nearly impossible to ignore the fighting going on around them. The emergence of fast moving *blitzkrieg* meant neutral defenses had to be completely re-thought. The air war pulled in neutrals if they did not adequately defend their airspace against any hostile invaders, even if they were only using it for transit. Irregular warfare also increased and dramatically eroded neutral sovereignty through fifth columns, spying operations, intelligence gathering, and secret plots.

The assumption inherent in the revisionist argument for any neutral is their conduct was somehow wrong. But it is difficult to reach this conclusion. How can we quantify the value of Sweden's contribution to Germany of ball bearings against the subjective worth of Allied signals intelligence or the rescue of the Danish Jews?

---

session, April 23, 1996; and United States Congress: Senate. Committee on banking, housing, and urban affairs, Swiss banks and attempts to recover assets belonging to the victims of the Holocaust hearing, 15 May 1997.

[6]Jan Heuman, "Final report on the Riksbank's wartime acquisition of gold" *Riksbank* (1997); International Commission of Experts, *Switzerland*: *national socialism and the Second World War* (2002), p. 245ff; Mugacia Commission, *Report of the commission of inquiry on gold transactions with the Third Reich* (1998).

[7]William Z. Slany "Preliminary study on US and allied efforts to recover and restore gold and other assets stolen or hidden by Germany during World War II" United States Department of State Report (May 1997).

[8]Ziegler, *The Swiss*, p.163.

[9]*Ibid*.

[10]Neff, Stephen C. *The rights and duties of neutrals*: *a general history* (2000, chapter 7).

If we limit this just to the economic perspective, the "code-of-conduct" group argues Sweden provided strategic goods to one or both parties which changed the course of the war in violation of the spirit, if not the letter, of the treaties. But these individual sales are frequently not placed in context.

For example, the Swedish provision of ball bearings and iron ore to Germany is taken out of perspective with the overall wartime trade situation and the Swedish need to earn German foreign exchange to buy food.[11] If we follow the logic of Mancur Olson, there are no strategic goods and therefore, individual items must be examined in context of overall trade.[12] The Swedes provided similar goods to the Allies throughout the war, even allowing Allied planes to land and boats to dock on Swedish territory to collect the much-needed Swedish exports.[13] These sales were, at least in part, made to obtain coal and petroleum products.[14] While it is beyond the scope of this inquiry, the question clearly arises: should the Swedish government have cut off trade with Germany and the Allies in order to satisfy their moral obligations, would it have been acceptable to allow its citizens to starve and freeze? Given the Swedish government is accountable to the Swedish people, the answer would seem to be "no." But this is ultimately a moral question. Although it does not directly address the moral issues, it places these arguments within the larger economic framework of wartime economic relations. Only then can we begin to understand them.

Several authors, most notably Neville Wylie and Christian Leitz, have approached this subject as a question of international relations. Wylie's edited volume *European Neutrals and Non-Belligerents during the Second World War* provides one of the most comprehensive currently available historical reviews of European neutrality during the Second World War period. This history includes both neutrals who were victims of aggression and those who survived the war.[15] However, Wylie defines the neutrals according to the period in which they perished, not the way in which they maintained their independence. The focus on the way they perished only tells half the story; importance has to be placed on determining why the countries who survived remained neutral throughout war.

In his comparative work, Leitz focuses on only the five European neutrals who survived the duration of the Second World War[16]; he reviews nominal German–

---

[11] Eric Golson, "Did Swedish ball bearings keep the Second World War going? Re-evaluating neutral Sweden's role" *Scandinavian Economic History Review*, 60:2 (June 2012), p. 165–182.

[12] Mancur Olson, *The economics of wartime shortage*: *history of British food supplies in the Napoleonic War and in World Wars I and II* (1965).

[13] Eric Golson, "Sweden as an occupied country? Swedish-Belligerent trade in the Second World War" in Jonas Scherner and Eugene N. White (eds.), *Hitler's war and Nazi Economic hegemony in occupied Europe* (2016).

[14] NA CAB122/241, correspondence dated 27 May 1942, November 1942 and between January to February 1943.

[15] Neville Wylie (ed.), *European neutrals and non-belligerents during the Second World War* (2002).

[16] Christian Leitz, *Nazi Germany and neutral Europe during the Second World War* (2000).

neutral trade statistics for a limited number of goods. Leitz asserts that each of the five neutrals in his study provided a significant contribution to the German war effort. But this results from a reliance on statistics from existing sources which focus on specific goods; many are based on political archives as opposed to more economically minded ones. As such, Leitz provides support for the revisionist argument by not placing this merchandise trade in a wider context. For example, the discussion of Sweden does not give us overall statistics for Swedish–German merchandise trade, but examples from what Leitz considers to be key goods, such as Swedish iron ore or Spanish wolfram.[17] In this way it becomes anecdotal: Leitz gives examples, but leaves his readers without a general understanding of the economic relationships between neutrals and belligerents. The present study builds on the work of both Wylie and Leitz by using standardized statistics which allow for cross-neutral comparison and reveal neutral preferences.

In three separate articles, Bruno Frey and various co-authors have used economic choices by bond investors to assess the perception of threats against belligerents and neutrals.[18] Using bond prices, Frey shows the turning points in the perceived threat against a neutral and which countries the markets believed would be successful in maintaining their independence through concessions. In particular, Frey is able to demonstrate that the price levels of Swiss bonds after June 1940 suggest the effectiveness of the Swiss concessions, with the markets perceiving little risk of invasion, especially relative to other neutrals.[19] An extension of Frey's work shows precursors to the Second World War trends. Values of neutral bonds decline significantly in the later years of the war when landings occur close to neutral territory (for example, the occupation of Russia by Allied troops in 1918). In the Second World War, we see more dramatic falls in the most threatened neutrals' bonds from the outbreak of war. So for both the neutrals' economic position and military strategy we look at the Second World War as the watershed moment.[20]

[17] *Ibid*., p. 71, 131.

[18] Bruno S. Frey and Marcel Kucher, "History as reflected in the capital markets: the case of World War II" *The Journal of Economic History*, 60:2 (June 2000), pp. 468–496; Bruno S. Frey and Marcel Kucher, "Wars and markets: how bond values reflect the Second World War," *Economica*, New Series, 68:271 (August 2001), pp. 317–333; Daniel Waldenström and Bruno S. Frey, "Using markets to measure pre-war threat assessments: the Nordic countries facing World War II," IMF Working Paper No. 676 (2006).

[19] Frey and Kucher, "History as reflected in the capital markets . . . ," pp. 484–485.

[20] S. Collet and E. Golson, "Neutral Central Bank financing in the Great War: how do neutral bond markets react to First World War financial and political events" Banque de France Research Seminar, November 2014.

Neutrality as Impartiality: Historical Context

As a legal concept, neutrality as impartiality first emerged in the early seventeenth century, when the first vague definitions of non-participation in war were penned by Hugo Grotius, a philosopher of modern natural law. Grotius argued: "From those who are at peace nothing should be taken except in case of extreme necessity and subject to the restoration of its value."[21] With reference to the duties of those who remain at peace, Grotius circumscribed neutral actions such that they should "show themselves impartial to either side in permitting transit, in furnishing supplies to his troops and in not assisting those under siege."[22] Grotius' version of neutrality as impartiality remains the most widely understood definition of the concept.

The Hague Conventions, Conference of London Declarations, Geneva Conventions, and other treaties of the late nineteenth and early twentieth century institutionalized the concept of neutrality as impartiality in war. The treaties detail the rights and duties of neutrals on land and at sea. Land rights include the inviolability of neutral territory, internment of combatants, rights to trade, and the neutrals' duties in protecting prisoners of war; naval rights include the ability to pass freely through the lines of war and maintain their trade relationships amidst war.[23] Legal scholar Stephan Neff argues these rights and duties were part of the requirements of a new form of impartiality, and in effect measurements of morality could be made from the neutrals' actions on these issues.[24]

Support for Neff's position can be found in the legal changes neutrality underwent after the First World War. New treaties established clear rules for the impartiality with which neutrality should function and restricted the behavior of neutrals in war. Declarations of neutrality had to meet the minimum standards of the Fifth Hague Convention, constraining the belligerents' actions against neutral states; a neutral's behavior was also restricted. With the adoption of these rules, the international community sought to lower the levels of uncertainty in wartime and prevent the spread of war to neutral countries.

The codification of neutrality as impartiality and agreements as to neutral rights provided neutral states with methods of redress against belligerents that violated their position. Examples include simple two-actor cases, such as the bombing by the Japanese of the neutral United States vessel *USS Panay* in a Chinese river in December 1937, for which the government of Japan paid slightly over $2.2 million

[21] Hugo Grotius, *De Jure Belli ac Pacis* [On the laws of war and peace], vol. 2:bk. 3 (1925), p. 783 (Translation).

[22] *Ibid*.

[23] See The Fifth Hague Convention of 1907, and the International Naval Conference of London, 1908–1909. For examples of the ways in which institutions lower systemic uncertainty, see Douglass North, *Institutions*, *institutional change*, *and economic performance* (1990).

[24] Neff, *The rights and duties of neutrals*.

dollars compensation in order to pacify the U.S.[25]; or the bombing of the Swiss city of Schaffhausen in 1944 by the U.S., for which compensation of $4 million was paid to Switzerland directly after the incident to maintain relations and Allied intelligence operations in the country.[26] These payments reinforce the concept of neutrality as impartiality according to Neff's criteria.

The Rise of Realist Neutrality

As previously suggested, the meaning of neutrality changed in the First World War. From this point, the concept of neutrality as impartiality, whether measured according to the Grotius or Neff versions, is altered. The military imbalance between neutrals and belligerents and particularly aggressive offensive tactics used in the Second World War confirm a version of neutrality which is based largely on economic *realpolitik*.[27] Neutral countries had rights and duties, but they were constrained by the power of the belligerent. What they were allowed to do was determined by their trade in material goods, services, capital, and labor as opposed to the law or moral considerations. Support for this position can be found in the writings of neutrality expert Nils Ørvik, who after the war posited:

> The outcome of the struggle to establish [a system of neutral rights] has at all times been entirely dependent on the economic and military strength, the strategic position and the perspicacity and persistence of the two sides. In short the products of two forces pulling in opposite directions, the final result being determined by the relative bargaining power of the parties.[28]

As Ørvik describes, with the increasing relative military strength of the Great Powers and their use of the cult of the offensive, neutrality evolved from impartiality into a realist system where the belligerents and the neutrals continually re-evaluated their relations on the basis of the power balance between them. In this realist power balance there were two critical components: the belligerent's strength and the neutral's deterrent against invasion. The belligerent's offensive strength over a neutral would have to be calculated with respect to military, economic, and political forces it could employ over the neutral and the credibility of such threats. Similarly, the neutral's defensive deterrent would involve some combination of the same forces, although the neutral not only had to ensure the credibility of its deterrent

[25]William Gerald Downey, Jr. "Claims for reparations and damages resulting from violation of neutral rights" *Law and Contemporary Problems*, 16:3 (Summer 1951), p. 488.

[26]Jonathan E. Helmreich, "The diplomacy of apology: U.S. bombings of Switzerland during World War II" *Air University Review* (May–June 1977); also by the same author, "The bombing of Zurich" *Aerospace Power Journal* (Summer 2000), pp. 48–55.

[27]Jack Snyder, "Civil–military relations and the cult of the offensive, 1914 to 1984" *International Security*, 9:1 (Summer 1984), pp. 108–146.

[28]Nils Ørvik, *The decline of neutrality 1914–1941* (1971), pp. 13–16.

force, but also broadcast its willingness to defend itself clearly to the belligerent. The rights and duties of neutrals were not important to the relative power balance; they were regularly ignored. The services the neutrals provided in the new version of neutrality outlined by Neff are important as part of this power balance; they are ultimately captured as economic payments in this chapter (they were not done through altruism).[29]

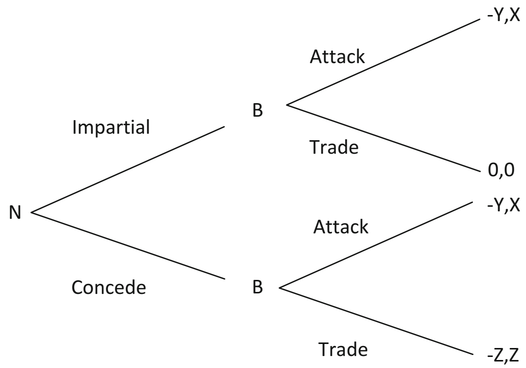
This chapter proposes a simple game in order to understand this relationship. This game is based on same principles as those introduced in "The Problem of Social Cost" by Ronald H. Coase.[30] In this realist power balance, there were two critical components: the belligerent's overall strength and the neutral's deterrent against invasion. It is also important to note the neutral's concession is constrained by the maximum potential cost the neutral might incur if invaded. The relative power balance between a single belligerent and a neutral can be defined according to the following relationship (Fig. 1):

This game from shows the estimation by the belligerent of deciding to wage war on the neutral. The belligerent will wage war if it thinks it will benefit more than it will lose; and the game shows how a neutral can, by offering concessions, make war by the belligerent less likely. X represents the net cost or benefit of invasion for the belligerent. It can be both positive (representing a benefit) and negative (representing a cost), depending on the relative positions of the neutral and the belligerent. For example, a shortage of military troops and equipment on the part of the belligerent would increase the costs of invasion; as would a strengthening of the neutral's deterrent. This increased cost of invasion would lower the benefits, and hence, reduce the value of X. These values are fixed for each instantaneous round of the game, but they can change with each new round. The game can be played until either the neutral is invaded or the neutral perceives threat from the belligerents to have ended.

Z represents the concession on the part of the neutral country. The neutral has the choice whether or not to provide concessions and, if it does provide them, the level and degree of the concessions provided. Concessions can include both political and economic means such as trade, labor, and capital provision. The value of Z is positive when neutrals provide concessions; but when they do not want to provide concessions and they adopt a strictly isolationist neutral position, Z is set equal to 0. It is also possible for Z to be negative, as in the examples of extortion from belligerents by neutrals profiting by the inherent advantages of their position.

[29]For more information on rational deterrence theory, see Christopher Achen and Duncan Snidal, "Rational deterrence theory and comparative case studies" *World Politics*, 41:2 (January 1989), pp. 143–169; Jack S. Levy "When do deterrent threats work?" *British Journal of Political Science*, 18:4 (October 1988), pp. 485–512; and John J. Mearsheimer, *Conventional deterrence* (1983); Richard Ned Lebow, "Rational Deterrence Theory: I Think, Therefore I Deter" *World Politics*, 41:2 (January 1989), pp.208–224; and Kenneth Waltz, *Theory of international politics* (1979).

[30]Ronald H. Coase, "The problem of social cost," *The Journal of Law and Economics*, 3:1960, pp. 144–171.



Where: B is the belligerent and N the neutral
X is the net benefit or cost of war to the belligerent
$X>0$ represents a benefit for the belligerent
$X<0$ represents a cost to the belligerent
Z is the cost of any concession on the part of the neutral country
$Z>0$ represents concessions (costs) for the neutral
$Z<0$ represents benefits for the neutral
Y is the cost to the neutral in the event of war ($Y>0$);
The cost to the neutral in the event of war must be greater than the concessions offered by the neutral ($Z>-Y$)
N chooses concessions if $Z<Y$
B chooses trade if $Z>X$
Concessionary trade is the equilibrium if $X<Z<Y$
There are indifference points at $X=0$ and $(X-Z)=0$

Fig. 1 Belligerent–neutral game: single belligerent model. Where: B is the belligerent and N the neutral. X is the net benefit or cost of war to the belligerent. $X > 0$ represents a benefit for the belligerent. $X < 0$ represents a cost to the belligerent. Z is the cost of any concession on the part of the neutral country. $Z > 0$ represents concessions (costs) for the neutral. $Z < 0$ represents benefits for the neutral. Y is the cost to the neutral in the event of war ($Y > 0$); The cost to the neutral in the event of war must be greater than the concessions offered by the neutral ($Z > -Y$). N chooses concessions if $Z < Y$. B chooses trade if $Z > X$. Concessionary trade is the equilibrium if $X < Z < Y$. There are indifference points at $X = 0$ and $(X - Z) = 0$

Examples include situations where neutrals have goods wanted or required by the belligerents and therefore, the belligerents surrender to them.

Belligerents' Perspective

For the belligerent power, the game hinges on the relative values of X and Z, namely, what it would win or lose in a war and what it would benefit in concessions. A costly invasion or increased military deterrence on the part of the neutral would lower the X value relative to both Z and 0. If the benefits of invading the neutral are net negative, as in $X < 0$, then it will not invade. But if the benefits are less than the concessions $Z>X$, then the belligerent power has no incentive to invade. The neutrals presented in this chapter are all clear cases of the value of the concessions exceeding the benefits of invasion such that $Z>X$; each example is with high costs of military invasion, and at least in the cases of Switzerland and Sweden, limited strategic benefits. Each provided specific and much demanded goods which were important for the belligerents' war efforts and which might have been lost in an invasion. If the benefits of war for the belligerent exceed the value of the concessions offered by the neutral $Z<X$, the belligerent power will invade. Neutrals including Belgium, Denmark, the Netherlands, and Norway are clear cases of large strategic benefits which sustained only limited military costs for Germany.

Neutrals' Perspective

This game has two players, insofar as the neutral power has the option to give concessions to lessen the belligerent's benefits from invasion. No neutral would logically give more than is necessary to deter invasion, so the neutral should always seek to minimize Z so that is only slightly larger than X. The game also hinges on the relative values of Y and Z. No neutral country will offer more in concessions than the costs of being invaded; therefore in the event of concessions in our model, $Z<Y$. Belgium, Denmark, the Netherlands, and Norway offered only limited concessions to the Germans in the period before they were invaded. The losses for these countries were also substantially less than others; Norway and Denmark even maintained civilian administrations for much of the war. The overwhelming German strategic return from invasion make it unlikely these neutrals could provide economic benefits sufficient to turn the balance away from invasion in the way neutral countries which were more costly to invade, such as Spain, Sweden, and Switzerland, could.

Realist Neutrality

The military disparity at the start of the Second World War was unusually stark. The costs of invasion for a belligerent were small compared to the benefits. The neutrals collapsed quickly in the face of German aggression. During the interwar period, the Great Powers made significant military advances, while the neutrals continued to rely on outdated technology and strategies. As the statistics in Table 1 suggest, these changes gave the Great Powers notable advantages.[31]

[31] Martin van Creveld, *Technology and war*: *from 2000 B.C. to the present* (1991); Larry H. Addington, *The patterns of war since the eighteenth century* (2004), pp. 176–189.

Table 1 Selected military strengths of European armies

Country	Men under arms	Tanks	Heavy guns	Air forces	Naval forces (home fleet)	Mechanized forces (divisions)
Belgium (1939)	600,000	10	Lacked heavy guns	50 modern	2 small warships	2 fully, 2 partially motorized
				200 obsolete		
Denmark (1940)	14,000	None	None	50 obsolete	2 coastal defense vessels	None
Ireland (1939)	7494	None	None	4 fighters	2 patrol boats	None
Netherlands (1940)	400,000	None	656 obsolete	50 obsolete	1 cruiser,	Bicycle corps
					1 destroyer	
					3 submarines	
					8 gun boats	
					5 minesweepers	
Norway (1939)	56,000	None	Few anti-tank and anti-air weapons	24 modern	7 destroyers	None
				44 older	1 minelayer	
				150 on order	40 smaller vessels	
Poland (1939)	1,000,000	300 medium tanks; 500 light tanks	4000 heavy field guns, of which 1154 'front line quality'	423 aircraft of which	4 destroyers	2 complete, 1 incomplete motor infantry; 110 armored cars
				16 modern bombers,	5 submarines	
				27 modern fighters	6 minesweepers	
Sweden (1937)	403,000	None	79 anti-aircraft	257 modern (1936)	47 assorted vessels	None
Switzerland (1939)	435,000	None	30–60 anti-aircraft guns	200 modern	Not applicable	Bicycle corps

(continued)

Table 1 (continued)

Country	Men under arms	Tanks	Heavy guns	Air forces	Naval forces (home fleet)	Mechanized forces (divisions)
France (1939)	5 million	3000 high quality	1927 guns	2200 planes, of which 1000 were modern	7 battleships	3 armored, 7 motorized infantry, 3 light mechanized; 3 motorized infantry
			7500 more on order		2 aircraft carriers	
					19 cruisers	
					58 destroyers	
					77 submarines	
United Kingdom (1939)	897,000	969 total	1900 total over 75 mm	7940 total, all types including old bi-planes	15 battleships	1 armored,
					4 aircraft carriers	4 motorized infantry
					37 cruisers	2 light mechanized
					105 destroyers	
					57 submarines	
Germany (1939)	3.74 million	4700 high quality	5000 total	2564 modern	5 battleships,	6 armored,
					1 heavy cruiser,	4 motorized infantry,
					4 light cruisers	4 light mechanized
					21 destroyers	2 airborne infantry
					57 submarines	

Source: Golson, "The Economics of Neutrality" (PhD diss., London School of Economics, 2011)

Against 16 mechanized German and French divisions, the best- equipped neutral country could boast only two mechanized and two partially mechanized infantry divisions. The Dutch and the Swiss still relied for mobility on bicycle corps and requisitioned civilian transportation. Others had no provision whatsoever for mechanized forces. Apart from 10 experimental Belgian models, the neutrals did not own a single tank, compared with some 4700 German models in 1939. The neutrals did have air forces, most notably the Swedes and Swiss, who had some 257 and 200 modern aircraft, respectively, but again these were small compared to the German force of 2564. Nor had the neutrals developed any new strategies to counter the new belligerent armies; instead, on 1 September 1939 they choose to deploy their infantry on the basis of strategies used in the Great War.[32]

It was to the advantage of the more powerful belligerents to invade and expropriate the forces of the militarily weaker neutrals. Within 9 months, the war had absorbed the Netherlands, which had a weak military and a strategy based on foreign intervention. The advantages of invading the Netherlands were particularly stark for the Germans, who could capture the costal ports, gold reserves, and manufacturing centers.[33] The strategic advantage of attacking France through Belgium made the latter's military defenses of little consequence from a strategic perspective. The limited economic cost of attacking compared with the benefits of not having to attack the Maginot line outweighed the few costs for the Germans. No economic concessions sufficient to counter these interests were deployed.

Less than 2 years into the war, only four continental European neutrals were left: Portugal, Spain, Sweden, and Switzerland. In the face of the military power of the belligerents, these remaining independent countries could not instantaneously reshape their military forces. A country such as Switzerland, now surrounded by Germany and similarly aligned Italy, could not instantly build a military force capable of countering German influence.[34] Nor could they build a Navy capable of protecting their merchant marine against the Allies. Instead they were dependent on economic concessions to counter the belligerents' desire to conquer them. Sufficient concessions have to be offered to assuage both parties without upsetting the other.

[32]Sweden relied on a 1925 strategy almost identical to its strategy in the Great War; see Ulf Olsson, "The state and industry in Swedish rearmament" in Martin Fritz et al, *The adaptable nation*: *essays in Swedish economy during the Second World War* (1983), p. 60. Switzerland relied on a defense-in-depth scheme from September 1939, later replaced by the national redoubt (réduit); see Henri Guisan, *Bericht an die Bundesversammlung Über den Aktivdienst 1939–1945* (1946), pp. 91–126.

[33]Bob Moore, "The Netherlands," in Wylie, *European Neutrals*, pp. 76–96.

[34]It took Germany 3.5 years (1935–1938) to build an air force of 50,000 men trained to use the various new types of aircraft. Defensive efforts were made, but these were insufficient against the superior German military of the time. Swiss General Guisan built a national redoubt (réduit) in July 1940, a series of linked defensive fortifications in the Alps based on the defeat of the German mechanized forces by the highly skilled Swiss troops. See Willi Gautschi, *General Henri Guisan*: *Commander in Chief of the Swiss Army in World War II* (2003), pp. 240–273. After 15 years of spending under $50 million altogether, Swedish military spending in 1939 increased to $322 million. It peaked in 1942 at $527 million (1938 dollars). Little information exists on Spanish military forces at this time.

The strength of this chapter can be found in the differences between the three countries. Geographically, economically, and politically, Spain, Sweden, and Switzerland were the neutral countries most affected by the Second World War. Sweden and Switzerland were physically surrounded by the Axis powers, cut-off from the rest of the world save for trade, capital flows, and communication unless mutually agreed to by both Germany and the Allied powers. In Spain, although less isolated, the dictator Francisco Franco's political desire to ally himself with the fascist cause meant Spain also faced a strict Allied blockade; all external Spanish trade required Allied consent. All three countries were threatened by the superiority of the belligerent forces.

The three countries also had very different economic starting points. Although Switzerland had the smallest population of the three, it was the richest, with per capita GDP of about $6390 in 1990 dollars. Economic development in Switzerland had been retarded in the 1930s by the Swiss government's adherence to the strong franc policy and the French-backed gold bloc.[35] It was until just before the war that things returned to pre-1930s levels.[36] Over the course of the war, the Swiss economy continued to contract slightly; as seen in Fig. 2, by 1943 the war retarded GDP growth by about 4 % of 1938 levels. Part of this GDP contraction could have been the result of changes in trade, which fell from 16 % of GDP before the war to 12 % by the end of the war. According to Maddison, there was a considerable 24 % increase in 1945 GDP as the war ended.[37] Swiss exports included finished goods and machinery. Imports consisted largely of primary products. Unemployment was less than 2 % and just over 10 % of the population under arms during the peak periods in 1939–1940.

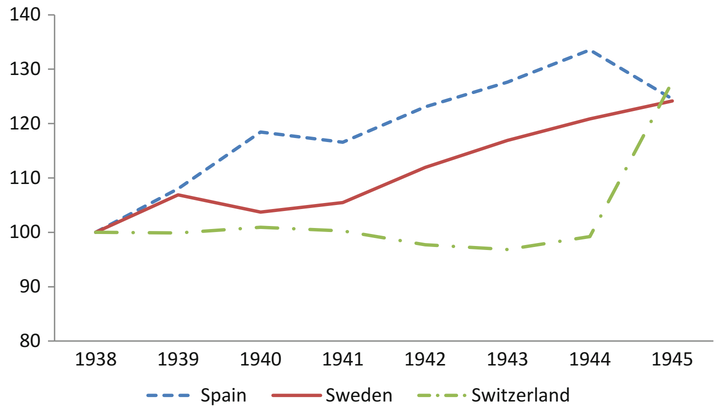


Fig. 2 Nominal GDP development in Spain, Sweden, and Switzerland, 1938–1945 (Source: Maddison, The World Economy, Table 1b)

[35]Maddison, *The World Economy*, Table 1b.

[36]*Ibid*.

[37]*Ibid*.

With vast sparsely inhabited lands, Sweden was somewhere in the middle of the extremes of Switzerland and Spain. Almost as big as Spain in territory, it had a population about one-and-a-half times the size of Switzerland with a slightly smaller overall GDP output. The result was a GDP per capita of about $4725, with about 16 % of the economy focused both on primary product and specialty goods exports at the start of the war.[38] Swedish GDP was stagnant for the first 3 years of the war, but then increased steadily starting in 1943; GDP increased 26 % from 1938 to 1945.

Spain had the largest population, economy, and territory of the three. But it was poor by European standards. Spanish GDP per capita was $1790 in 1990 dollars; amongst the main European belligerent powers, only Italy had worse GDP per capita figures. Economic development in Spain had been severely reversed by the Spanish civil war, with overall GDP 29 % below the 1928 peak of $63.57 billion and the 1935 pre-war figure of $63.48 billion.[39] As a result, going into the war Spain had significant spare productive capacity which would allow for a rapid increase in output. But during the Second World War, the Spanish economy increased output by some 33 % between 1939 and 1944. Main exports throughout the war were largely primary products, including wolfram, furs, and food. Unemployment was rife, and although no reliable statistics are available, it is suggested to have been as high as 40 % of the population during this period; although it is important to note skilled labor was in short supply throughout the wartime period.[40]

Whether it is through physical encirclement, changes in trade patterns, or political rhetoric, Spain, Sweden, and Switzerland were all threatened by the war. Obviously the main aggressors were physically, militarily, and economically larger than these three neutrals. Both the UK and Germany were wealthier than Spain or Sweden on a per capita basis ($5983 and $5126, respectively)[41]; with the territories under belligerent control came access to raw materials, skilled labor, and in some cases, capital. While there can be no question of consistent shortages of all three throughout the war for the belligerents, their particular policies of rationing material goods often made these shortages worse in neutral countries. Despite some political overtures, each neutral was forced to engage in economic *realpolitik* in the areas of goods trade, capital and skilled labor in order to ensure their survival. So while this section has shown these three had different starting points and unique political perspectives, this chapter ultimately observes similar economic patterns for the three countries and demonstrates they survived the war by analogous means.

---

[38] *Ibid*.

[39] *Ibid*.

[40] Enrique Prieto Tejeiro, *Agricultura y Atraso en la España Contemplo-Ranea* (1988), pp. 58–59.

[41] Angus Maddison, *The World Economy*: *Historical Statistics* (OECD, 2003), Table 1a–1d; Mark Harrison (ed.), *The Economics of World War II* (1998), p. 3.

Economic Concessions

Spain, Sweden, and Switzerland survived their encirclement by Germany through carefully conceived economic deterrents and concessions. Despite the different political origins of the three neutrals' policies, their economic concessions turn out to be quite similar. Following the existing literature on these countries wartime activities and Mancur Olson's focus on physical goods in wartime situations, this chapter now turns to an analysis of goods and monetary flows in the pattern of economic concession. When Hitler wielded most power, the Germans benefited from increased trade volumes, favorable pricing and capital flows. This favoritism towards Germany was particularly notable during its period of strength after the defeat of France in June 1940 and declined following the increasing Allied successes after 1943. The Germans benefited from increased trade volumes and favorable pricing during the apex of Hitler's power. Switzerland, surrounded by the Germans, provided sizeable trade credits. But German control is never absolute.

The two most isolated countries, Switzerland and Sweden, also provided concessions to the Allies in order to maintain the recognition of their neutrality and access to Allied markets. Small goods critical to the Allies' war effort were smuggled through every available channel around the blockades; services were extended to the Allies. Spain provided pricing benefits to the countries with the greatest relative power and which could provide it with goods that it needed and wanted—these countries being the United Kingdom and Germany. But the common thread among the three is they did not excessively defer to one belligerent or another, but rather like any country, they tried to obtain the best possible terms for themselves in their relations with the belligerents. They did not de facto become part of the Axis or Allied groups.

Amongst the tactics used, the Swedish government ensured both parties received what they desired and prevented the other from objecting using information asymmetries. The normal trade agreements system clearly defined the quantities of particular exports. The Swedish government disregarded the agreed terms of its annual war trade agreements with the belligerents, especially for critical goods, in order to fully satisfy both belligerent groups.[42] Quantities of goods, payment terms, and other aspects were altered according to wartime conditions. The Swedes provided the Germans with the iron ore, ball bearings, and machine tools critical for the German war effort; in a clear concession to the Germans, 1941, the peak year of German military power, was also the peak year of Swedish exports to Germany. The Swedes benefited from nominal goods import surpluses with both belligerent groups and were able to maintain favorable prices in German–Swedish trade throughout the period. The Swedes did not have to provide trade credits to the Germans. In the case of German–Swedish trade, the excess merchandise imports favoring Sweden

[42]Eric Golson, "Sweden as an Occupied Country? Swedish-Belligerent Trade in the Second World War" in Jonas Scherner and Eugene N. White (eds.), *Hitler's war and Nazi economic hegemony in occupied Europe* (2016).

amounted to 1.8 billion Kronor, equaling 14.9 % of Swedish 1938 GDP.[43] Despite the large goods surplus, it is evident that the Swedish government acted within the previously described realist neutral framework.

In order to maintain its links outside the German trade bloc, the Swedes offered concessionary prices to the Allies and permitted the transportation of ball bearings, machine tools, and other items important for the Allied war effort, documented here for the first time. These items were transported by air and sea through the blockade from Sweden to the UK, with trade increasing steadily until 1945. As part of these activities, the Swedish allowed Allied quasi-military transports to use Swedish territory for blockade running, a concession which was otherwise incompatible with international neutral law.[44] The Swedish government was also willing to suspend exports to Germany in October 1944, to benefit the Allies given that the Swedes were still surrounded by Axis powers. When contrasted against the pro-German actions, this clearly shows the Swedish government was acting pragmatically throughout the war, giving concessions where necessary to maintain Sweden's independence.

Despite being engulfed by the German bloc and its customs regime, the Swiss government was also able to maintain trade with the Allies. The Allies exerted financial and diplomatic pressure on the Swiss government. After Britain stated that its test for Swiss independence was the maintenance of an export surplus, the Swiss government provided the Allies with relative price benefits and a large export surplus (in nominal and real terms) and ensured that export trade volumes were maintained.[45] In what was a particularly large concession to the Allies, exports included much-needed military goods, such as watches and guns.

To satisfy German military pressure, the Swiss offered a clearing deficit which rose to 990 million Swiss Francs, equivalent to 12.0 % of Swiss 1938 GDP between 1941 and late 1943.[46] In order to avoid trade retaliation, the Swiss also exploited information asymmetries and tried to prevent the disclosure of these sums to the Allies. Switzerland also provided the Germans with increased quantities of specialty exports; price increases were not as notable as in the Swedish or Spanish cases, despite the fact most of the goods were sold on credit.[47] Although Switzerland was still surrounded by the Germans until February 1945, after mid-1943, the Germans were at an increasing disadvantage, in both net trade and pricing terms.

The Spanish–belligerent relationship also fits within the suggested neutral–belligerent framework. The Spanish government was initially in a weak position against the German government and therefore offered concessions on exports.

[43]Eric Golson, "German and British balance of payments with European neutrals in the Second World War" forthcoming; Maddison, *The World Economy*, Table 1b.

[44]Golson, "Swedish ball bearings".

[45]NA FO837/960-962.

[46]Maddison, *The world economy*, Table 1b.

[47]Eric Golson, "Swiss trade with the allied and axis powers during the Second World War" *Jahrbuch Für Wirtschaftsgeschichte*, 2014:2 (November 2014), p. 71–97.

However, the Spanish supplied the dense metal wolfram and the Spanish government was soon able to extract increasingly stiff price concessions from all of the belligerents using the *quid pro quo* trading system. It gained substantial technical and military equipment from the Germans which could have otherwise been used in the German war effort. The Spanish quid pro quo trading system allowed the Spanish government to best the Allies forcing them to pay for Spanish exports in gold, despite the Allied control of Spain's energy supplies. Moreover, the Spanish maintained the Allied–Axis competition right until the Allied armies reached the Spanish border. As a result, Spain's wartime trade position is consequently best characterized as one of pragmatic self-benefit, not wartime domination by Germany. The seemingly close political alliance between Hitler and Franco clearly had limited impact on the economic relationship of the two countries, especially during the later years of the war.

Whether in trade or capital flows, the three neutrals in this study survived the war (at least economically) through carefully designed concessions and deterrents. With military deterrents alone, it would have been very difficult for these neutrals to resist German or Allied aggression. In order to maintain a realist power balance between a neutral and a belligerent, the neutral was compelled to use economic means for survival as explained by the multi-player game provided in the introduction. This chapter has demonstrated that in this context modern neutrality cannot be impartiality, with sales of identical quantities of goods and the provision of equal rights. Rather, as the neutrals trade and capital concessions in the Second World War show, neutral self-preservation is an exercise in economic realism, where the game necessitates changing the level of economic concessions from the neutral or benefits to the neutral to correct for the relative power balance between a neutral and a belligerent at any given time.

Archival and Published Primary Sources

United States Congress: Senate. Committee on Banking, Housing, and Urban Affairs, Swiss banks and the status of assets of Holocaust survivors or heirs: hearing before the Committee on Banking, Housing, and Urban Affairs, United States Senate, One Hundred Fourth Congress, second session, April 23, 1996 (Washington, 1996)

United States Congress: Senate. Committee on Banking, Housing, and Urban Affairs, Swiss banks and attempts to recover assets belonging to the victims of the Holocaust hearing before the Committee on Banking, Housing, and Urban Affairs, United States Senate, One Hundred Fifth Congress, first session on the inquiry into the assets of Holocaust victims deposited in Swiss banks and the issues surrounding the recovery and restoration of gold and other assets looted by Nazi Germany during World War II, and the acts of restitution which must follow, Thursday, May 15, 1997 (Washington DC, 1997)

National Archives, Kew Gardens, London, UK (NA)

NA CAB122/241

NA FO837/960-962

References

Aarons, Mark, and John Loftus. 1998. *Unholy trinity: The Vatican, the Nazis, and the Swiss Banks*. New York: St. Martin's Griffin.

Achen, Christopher, and Duncan Snidal. 1989. Rational deterrence theory and comparative case studies. *World Politics* 41(2): 143–169.

Bower, Tom. 1997. *Blood money: The Swiss, the Nazis and the looted billions*. London: Pan Books.

Churchill, Winston. 1953. *The second world war: Triumph and tragedy*, vol. 6. Boston: Houghton Mifflin.

Coase, Ronald H. 1960. The problem of social cost. *The Journal of Law and Economics* 3: 144–171.

Downey Jr., William Gerald. 1951. Claims for reparations and damages resulting from violation of neutral rights. *Law and Contemporary Problems* 16: 3.

Frey, Bruno S., and Marcel Kucher. 2000. History as reflected in the capital markets: The case of world war II. *The Journal of Economic History* 60(2): 468–496.

Frey, Bruno S., and Marcel Kucher. 2001. Wars and markets: How bond values reflect the Second World War. *Economica, New Series* 68(271): 317–333.

Fritz, Martin. 1982. *The adaptable nation: Essays in the Swedish economy during the Second World War*. Stockholm: Almqvist and Wiksell International.

Golson, Eric. 2012. Did Swedish ball bearings keep the Second World War going? Re-evaluating neutral Sweden's role. *Scandinavian Economic History Review* 60(2): 165–182.

Golson, Eric. 2014. Swiss trade with the allied and axis powers during the Second World War. *Jahrbuch Für Wirtschaftsgeschichte* 2: 71–97.

Golson, Eric. 2016. Sweden as an Occupied Country? Swedish-belligerent trade in the second world war. In *Hitler's war and Nazi economic hegemony in occupied Europe,* ed. Jonas Scherner and Eugene N. White. Cambridge: Cambridge University Press.

Grotius, Hugo. 1925. *De Jure Belli ac Pacis* [On the laws of war and peace]. London: Humphrey Milford. 2:3.

Guisan, Henri. 1946. *Bericht an die Bundesversammlung Über den Aktivdienst 1939–1945* [Report to the Federal Council about (my) Active Service 1939–1945]. Lausanne.

Hanhimaki, Jussi M. 2016. Non-aligned to what? In *Neutrality and neutralism in the global cold war: Between or within the blocs?* ed. Hanhimaki Bott. New York: Routledge.

Harrison, Mark. 1998. *The economics of world war II: Six great powers in international comparison.* Cambridge: Cambridge University Press.

Helmreich, Jonathan E. 1977. The diplomacy of apology: U.S. bombings of Switzerland during World War II" *Air University Review* May–June.

Heuman, Jan. 1997. *Final report on the Riksbank's wartime acquisition of gold*. Stockholm: Riksbank.

International Commission of Experts. 2002. *Switzerland: National socialism and the second world war* Zurich: Pendo.

Jordan Ends Neutrality, Assailing Allied War Effort. *The New York Times*, February 7, 1991.

LeBor, Adam. 1997. *Hitler's secret bankers* Secaucus: Carol Pub. Group.

Lebow, Richard Ned. 1989. Rational deterrence theory: I think, therefore i deter. *World Politics* 41(2): 208–224.

Leitz, Christian. 2000. *Nazi Germany and neutral Europe during the Second World War.* Manchester: Manchester University Press.

Levy, Jack S. 1988. When do deterrent threats work? *British Journal of Political Science* 18(4): 485–512.

Maddison, Angus. 2003. *The world economy: Historical statistics.* Paris: OECD.

Mearsheimer, John J. 1983. *Conventional deterrence*. Ithaca: Cornell University Press.

Mugacia Commission. 1998. *Report of the commission of inquiry on gold transactions with the Third Reich.* Madrid.

Neff, Stephen C. 2000. *The rights and duties of neutrals: A general history*. Manchester: Manchester University Press.

North, Douglass. 1990. *Institutions, institutional change, and economic performance.* Cambridge: Cambridge University Press.

Olson, Mancur. 1963. *The economics of wartime shortage: A history of food supplies in the Napoleonic war and in world wars I and II*. Durham: Durham University Press.

Ørvik, Nils. 1971. *The decline of neutrality, 1914–1941*. London: Frank Cass & Co.

Rings, Warner. 1990. *Raubgold aus Deutschland: die Golddrehscheibe Schweiz in Zweiten Weltkrieg.* Ascona.

Sarasin, Philipp and Regina Wecker. 1998. *Raubgold Réduit Flüchtlinge: Zur Geschichte der Schweiz im Zweiten* Weltkrieg. Zurich.

Slany, William Z. 1998. "United States Interagency Report on U.S. and Allied Wartime and Postwar Relations and Negotiations with Argentina, Portugal, Spain, Sweden and Turkey on Looted Gold and German External Assets. Washington, DC.

Snyder, Jack. 1984. Civil-military relations and the cult of the offensive, 1914 and 1984. *International Security* 9(1): 108–146.

van Creveld, Martin. 1991. *Technology and war: From 2000 B.C. to the present.* London.

Vincent, Isabel. 1997. *Swiss Banks, Nazi Gold, and the Pursuit of Justice.* New York.

Waldenström, Daniel, and Bruno S. Frey. 2006. Using markets to measure pre-war threat assessments: The nordic countries facing world war II. IMF Working Paper No. 676. Zurich: Institute for Empirical Research in Economics.

Waltz, Kenneth. 1979. *Theory of international politics.* Reading: Addison-Wesley

Wylie, Neville (ed.). 2002. *European neutrals and non-belligerents during the Second World War*. Cambridge: Cambridge University Press.

Ziegler, Jean. 1998. *The Swiss, the Gold, and the Dead: How Swiss Bankers Helped Finance the Nazi War Machine*, Trans. John Brownjohn. New York.

Index

J. Eloranta et al. (eds.), *Economic History of Warfare and State Formation*, Studies in Economic History, DOI 10.1007/978-981-10-1605-9

V

W

Z