G

Galbraith, John Kenneth

Alexandre Chirat
Triangle, Université Lumière Lyon II, Lyon, France

Abstract

John Kenneth Galbraith was a post-Keynesian and an institutionalist economist. He is famous for having built an unconventional vision of American Capitalism in his postwar trilogy (1952b, 1958, 1967). His socioeconomic analysis deals with subjects such as theory of the firm, managerial revolution, public spending, theory of consumption, price control (1952a), financial crises (1955), power relationship (1983), and more generally with the dynamic of capitalist institutions.

Biography

"Ken" Galbraith (October 15, 1908–April 29, 2006), born in Ontario, was trained after the 1929 Great Crash and passed a Ph.D. at Berkeley in Agricultural Economics (Galbraith 1981). He moved to Harvard during the 1930s and spent a year in Cambridge among Keynes' fellows (UK). During World War II, Galbraith led the Office for Price Administration. He also oversaw a committee for the evaluation of the effects of American bombings on the German War Economy. Galbraith then made a journalist career at *Fortune Magazine*. At the political level, he was a leading figure, with his friend Arthur Schlesinger, of the *American for Democratic Action* (*ADA*) and a special adviser of Democratic candidates running for the Presidency. In 1958, he lamented the "private opulence and public squalor" in his best-seller *The Affluent Society* (1958). J.F. Kennedy named him Ambassador to India in 1961. This experience delayed the publication of his masterwork, *The New Industrial State*, which gave rise to several debates among economists. He spent the rest of his academic life struggling against neoclassical economics and its "conventional wisdom," against the fact that economists tended to overlook real social issues in favor of purely analytical constructions and intellectual puzzle. For these reasons, John Kenneth Galbraith was one of the most renowned public intellectuals of postwar Anglo-Saxons societies. Indeed he always relied on public enlightenment and developed his ideas for large audiences. For instance, he took part in a TV series entitled *The Age of Uncertainty*, broadcast on the BBC during the 1970s, which aimed to get people acquainted with economic knowledge (Parker 2005).

Innovative and Original Aspects

In his work Galbraith has always insisted on the idea that power *de facto* tends to overthrow *de jure* relationship. For instance, in line with Berle and

© Springer Science+Business Media, LLC, part of Springer Nature 2019
A. Marciano, G. B. Ramello (eds.), *Encyclopedia of Law and Economics*,
https://doi.org/10.1007/978-1-4614-7753-2

Means (1991), he explained that corporations are ruled by those who possess and control crucial factors of production. In the "planning system," this factor is "organized intelligence" (1967). That is why he thought that a "technostructure" rather than its owners controls modern corporations.

His contribution to Law and Economics is linked to his analysis of the structure of the American economy. In the first opus of his trilogy, *American Capitalism*, he explained that monopoly and oligopoly are now the norm rather than the exception. Thus, the Antitrust Legislation, which is embodied by the Sherman and Clayton Acts, is irrelevant. Whereas some economists pleaded for an enforcement of the antitrust laws against increasing monopoly and monopolistic power, Galbraith explained that competition among giants (corporations, unions, workers' and buyers' cooperatives) is susceptible to lead to some equilibrium (1952b). This is at the heart of his idea that "countervailing power" is the new regulative mechanism of the American economy. "In Galbraith's theory the offsetting power of groups on opposite sides of the market fills the breach, and market power is thereby kept from upsetting the allocative and distributive mechanism" (Peterson 1957). In this regard, he dismissed the Clayton Act for serving as a means to dismantle the bargaining power of Unions rather than the original market power of giant modern corporations.

In *The Affluent Society*, Galbraith explains that antitrust laws were defended by economists since they aim to protect consumers from abuses and to struggle against inequalities. However, in the affluent society, "increased production" is becoming a substitute to "redistribution" (1958). Thus, antitrust laws are anachronistic in American modern capitalism. They are a "charade" according to Galbraith since he considers that big corporations are more efficient to achieve increasing outputs. He notes that "it is at least amusing that during both World Wars the enforcement of the antitrust laws, the traditional design for ensuring greater resource-use efficiency, was suspended" (1958). With the second opus of his trilogy, Galbraith belongs to "a minority of economists," who consider that "the antitrust laws have outlived their usefulness" and are unable "to deal with the phenomenon of oligopoly" (Breit and Elzinga 1968).

Galbraith pursues such a line of reasoning in *The New Industrial State*. American conception of antitrust laws is a paradox. On the one hand, there is a deep ideal of competition requiring the absence of market power, thanks to the enforcement of the antitrust laws. On the other, there is an ideal American way of life, which requires big corporations in order to achieve increasing outputs. "No one can ask [a firm] to be an oligopolist for the purposes of capital investment, organization and technology and to be small and competitive for the purposes of prices and allocative efficiency. There is a unity in social phenomena which must be respected" (1967). There are some other contradictions. Regarding antitrust legislation, monopoly is illegal. But oligopoly without overt collusion, even if it has the same economic consequences, is not. Moreover, antitrust laws are sometimes used to condemn two small merging competitors rather than a single firm that controls a greater extent of market shares. His judgment is without appeal: "present enforcement [of antitrust laws] attacks the symbol of market power and leaves the substance" (1967).

Instead of relying on antitrust laws, Galbraith thinks that direct social control of corporations is more efficient. This preference is first understandable because he thinks as every institutionalist that concentration is "the inevitable product of economic imperatives" (Muller 1975). Then, his experience at the head of the Office for Price Administration convinced him that direct negotiation between corporations and public administrations is workable. He argues that price control is a more effective means for avoiding oligopolies' abuses (1952a, 1967). Muller, considering the relevance of Galbraith's analysis, adds that direct public control and indirect ones, through antitrust laws, could be complementary so as to reduce the undesirable effects of market power.

This succinct summary of Galbraith's conception of antitrust laws is one of his numerous innovative analyses. Nowadays such a reflexion appears even more original since Competition Policy, especially in the European Union, has

become a substitute rather than a complement to Industrial Policy.

Impact and Legacy

John Kenneth Galbraith did not give birth to a school of thought, since on the one hand he was on the fringe of academia and on the other hand he did not like to teach. James Galbraith sums up that "the Galbraith paradox is that the great theorist of organizations worked alone – he was an intellectual entrepreneur" (2007). Some of his works are largely known, others deserve to be. The impact of Galbraith's works on postwar economics was however crucial, since he always refused to abandon his critical glance towards economics.

Cross-References

► Institutional Economics

References

Berle A, Means G (1991) Modern corporation and private property. Transaction Publisher, New Brunswick
Breit W, Elzinga K (1968) Oligopoly, the Sherman Act and "the new industrial state". Soc Sci Q 49:49–57
Galbraith JK (1952a) A theory of price control. Harvard University Press, Cambridge, MA
Galbraith JK (1952b) American capitalism. Houghton Mifflin Company, Boston
Galbraith JK (1955) The great crash. Houghton Mifflin Company, Boston
Galbraith JK (1958) The affluent society. Houghton Mifflin Company, Boston
Galbraith JK (1967) The new industrial state. Houghton Mifflin Company, Boston
Galbraith JK (1981) A life in our time. Houghton Mifflin Company, Boston
Galbraith JK (1983) Anatomy of power. Houghton Mifflin Company, Boston
Galbraith JK (2007) Foreword to the new industrial state, 4th edn. Princeton University, Princeton
Muller WF (1975) Antitrust in a planned economy: an anachronism or an essential complement? J Econ Issues 9:159–179
Parker R (2005) John Kenneth Galbraith, His Life, His Politics, His Economics. University of Chicago Press
Peterson S (1957) Antitrust and the classic model. Am Econ Rev 47:60–78

Games

Ulrich Blum
Faculty of Law and Economics, Martin-Luther-University of Halle Wittenberg, Halle, Germany
University of International Business and Economics, Beijing, China

Abstract

This entry begins by setting games within the general framework to demonstrate their conceptual importance for describing this activity of humankind. Then the concept of merging economics and mathematics into what is called game theory is developed. Descriptions of important properties, definitions, and typical examples of games are provided.

Definition

A game is a structured interaction of at least two players. It consists of rules that set the institutional framework of possible moves. These lead to outcomes that are subject to individual preferences, which generate the incentives to play. If individual ends are incompatible, dilemmata develop, which make the underlying strategic interaction particularly visible.

Games and Science

Even before John von Neumann and Oskar Morgenstern (1944) merged mathematical and economic theory in their groundbreaking work *Theory of Games and Economic Behavior* games were attracting interest in other fields. For instance, Johan Huizinga (1933) looked at games from a psychological perspective showing that they are a fundamental force and a formative element of culture as defined by philosophical anthropology. Games date back to a time before the development of humankind: Many creatures, especially primates, but also other mammals as

well, show game-related behavioral patterns as an expression of rivalry. From a phylogenetic perspective, animals play, and from an ontogenetic perspective, games are incorporated into the existence and the development of the personal individuality of humankind. Games represent a field of tension between conflict – mankind's aggression – and cooperation, the attempt to bind interests. If we say "a game is captivating," then we metaphorically combine these antagonistic aspects in one sentence. In history and social sciences, managing conflict and mapping it in parlor games have always accompanied human development. In fact, John von Neumann (1928) looked at this issue in his first book on game theory. The greatest works on strategy, which are often compared today to game-theoretic models, were developed by scholars like Sun Zi (544–466 BC) from China or Carl von Clausewitz (1780–1831) from Germany. The *Art of War* (~500 BC, 2003, 2007) and *On War* (*Vom Kriege*, 1832) are two classic books about leadership under conditions that involve strategic, operational, or tactical interaction – an aspect of humankind that is also found in important religious works such as the Talmud, the Bible, and the Koran.

The work of John von Neumann (1928) and Oskar Morgenstern triggered not only a broad scientific development of game theory, such as the work by John Nash (1950) on equilibria published half a decade later, it also triggered a very strict observation of reality and its dilemma situations. On the level of political thinking, the concept of zero-sum games is perhaps the best-known outflow and potentially the least understood. In terms of the military, the concept of a credible threat during the Cold War was the direct result of game-theoretic thinking.

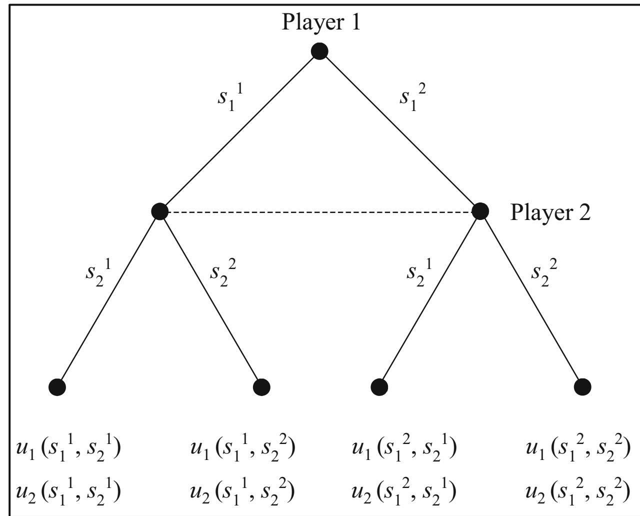
Economic and Political Aspects of Game Theory

If human rivalry is to be oriented toward a common good, and if the social contract written by societies is to be productive, the distinction between institutional rules and the moves allowed within this framework is of utmost importance. This relates to extending competition in markets to the idea of competition among institutions, especially public institutions. According to the philosophy of Immanuel Kant (1787) and, in particular, his concept of the "categorical imperative," economic institutions should be designed in such a way that the maxim of their implementation can be accepted as a general rule or regulation. Thus, dilemmata reach beyond individual rivalry, extending to the interaction between institution setters, institutions, and their efficiency in relation to other institutions, thus binding together economics and law. This lies at the core of the very successful German social market economy in the tradition of Walter Eucken (1952) and Ludwig Erhard (1957). Rules are to be set so that they are self-enforcing and follow mankind's open incentive concept. In such a world, games have two properties: rivalry over better arrangements at the higher institutional level and rivalry within the economic or social world below this roof. This addresses two aspects accentuated by game theory: hierarchy and dynamics. Who and how are the rules set for games, how open are they to long-term change, and who is in charge? Furthermore, with regard to moves (the complete set of which is called strategies), how can players learn to distinguish between "good" moves and "bad" moves? The final chapter will take up this subject using institutional competition as an example.

The rest of the entry is structured as follows: First we will look at the information side of game theory and the prerequisites for efficient activities. Then we will look at how one type of game may switch to other type when circumstances change. Finally, we will look at dynamic contexts.

The General Structure of Games

A game is a model of strategic (or tactical) interaction between players who have sets of possible actions. Since the model captures this interaction between the players, outcomes are interdependent. Dixit and Nalebuff (2010) is an introductory text that requires little formal experience. Standard textbooks are Rasmusen (2006) or Osborne and Rubinstein (1994). Binmore (1991) adds humor



Games, Fig. 1 Extensive form of a two-player game

Games, Fig. 2 Prisoner's dilemma game

		Player 2	
		silence	fink
Player 1	silence	– 1 / – 1 A	C – 9 / 0
	fink	0 / – 9 B	D – 6 / – 6

to theory. Players come up with strategies *s* and have preferences that are modeled by ordinal payoffs or cardinal functions, for instance, production functions *p(.)* or utility functions *u(.)*. In its extensive form, the game starts with two possible moves from the first player's strategy set ($s_1^{\ 1}$ and $s_1^{\ 2}$). This is followed by the moves of the second player. Figure 1 illustrates the sequential structure.

One of the best-known games is the prisoner's dilemma. The fundamental idea is straightforward: A murder has been committed and the police apprehend two suspects tramping in the area. They lock them up separately and have them interrogated by an attorney. The vagrants would be best off if they both remained silent; then they would be sentenced to only 1 year each for vagrancy. However, if one of them finks (cheats, defects) and accuses the other, he will be set free and the other will be jailed for 9 years. If both accuse the other person, they will both be jailed for 6 years. To all extent and purposes, it would be socially efficient to remain silent. However, under the absence of coordination, the best possible strategy is to accuse the other, which, however, produces an inferior outcome (Fig. 2).

If both defendants remain silent, the outcome is better than if they accuse each other (both payoffs are strictly higher values). The best strategies for one player against the other are underlined. However, this advantage does not exist with respect to the other two pairs of payoffs. In addition, the payoff resulting from mutual accusation remains unchanged in the sense that neither prisoner can deviate without disadvantaging himself. This leads to the following definition:

> A set of player strategies is a Nash equilibrium if no deviation of any individual player's strategies (i.e. given the strategies of the other players) results in a better outcome for that player.

Results in field D are thus underlined twice as they are the Nash equilibrium.

Games, Fig. 3 Tragedy of the commons game

		all other players				
		cooperate	defect	act at 10% in consort	act at 30% in consort	act at 60% in consort
player A	cooperate	10/10	2/12	2.8/11.8	4.4/11.4	6.8/10.8
	defect	12/2	3/3	11.1/2.1	9.3/2.3	6.6/2.6

A C
B D

One of the most interesting games in the genre of prisoner's dilemmas is the tragedy of the commons (Hardin 1968). In this game, a joint resource is distributed. An overexploitation and destruction of the resource can only be prevented if all parties agree to the terms of exploitation. However, under the absence of coordination, each individual will seek to achieve his/her own optimal outcome, and the resource will break down. The resource-efficient outcome will only be achieved if there is a high enough expectancy to cooperate. Figure 3 shows that this is the case if 60% act in consort. If player A then decides to defect, he will be worse off (6.6 instead of 6.8 points). This is not yet the case when only 30% act in consort.

There are many interesting dilemma situations: The chicken game models a situation where two cars approach each other on a narrow road with the question being, which car gives way to which? The battle of the sexes models a situation where a couple wants to spend their weekend away from home; he wants to go to a soccer game, while she wants to go to the theater. Both strategies on their own are incompatible with one another, and there is no joint utility-maximizing outcome. The problem in this game is that mixed strategies are not possible. A compromise can only be found, and socially efficient results can only emerge, if a mix between the two is possible, for instance, if two weekends are taken into consideration.

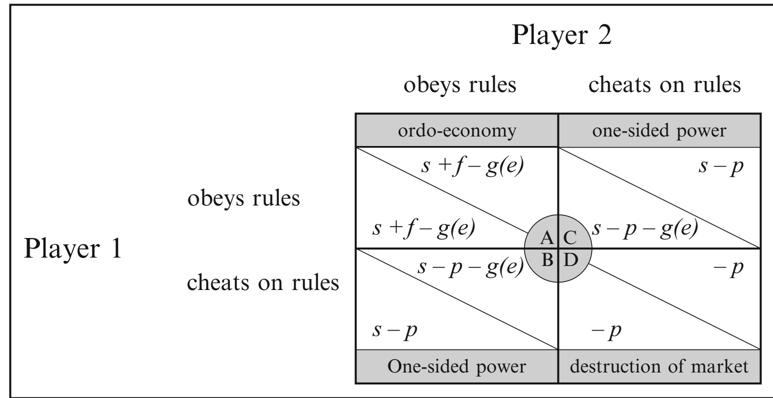
From these discussions, two other notions can easily be derived:

- In zero-sum games, the extent of the loss on the one side is identical to the gain on the other side.
- Zero-sum games are a special variant of constant sum games.
- In policy discussions, people think situations are zero-sum situations even when they truly are not. For instance, following Ricardo's classic theory of comparative costs, trade is a redistribution that makes both parties better off.

Games with Functions as Payoffs

Often, games are set within a dynamic context as they develop in an evolutionary way or are played repetitively. Players are then able to gain experience, which may encourage them not to fall into inefficient Nash equilibria, e.g., into the prisoner's dilemma trap. Another aspect is the changing of payoffs in dynamic utility or production functions, as in the following case which combines the economics of competition with the legal side of good regulations and governance. The example follows (Blum et al. 2005, section 7.2.2). In a competitive environment, the antitrust organization sets penalties, p, if competition rules are broken. The standard reward from competition is

Games, Fig. 4 Competition game



G

Games, Table 1 Basic game structures

Basic game	Line player (player 1)
Prisoner's dilemma	$B_1 > A_1 > D_1 > C_1$
Chicken game	$B_1 > A_1 > C_1 > D_1$
Assurance game	$A_1 > D_1 > C_1$ and $A_1 > B_1$
Social optimum	$A_1 > C_1 > D_1$ and $A_1 > B_1$

$s \geq 0$; it becomes zero if all players violate competition laws. An extra externality or fringe profit, $f \geq 0$, emerges if all players comply with the rules. The effort of players is given as $e \geq 0$ and the related disutility is $g(e) \geq 0$ (Fig. 4).

Table 1 lists three types of games and the structure of their payoffs, to which we have added a social-optimum outcome.

If we evaluate the outcomes, we obtain:

- **Prisoner's dilemma:** (player 1, $B_1 > A_1 > D_1 > C_1$; player 2, $C_2 > A_2 > D_2 > B_2$)

$$B_1 > A_1 (\text{or } C_2 > A_2 \text{ resp.}) \Leftrightarrow s - p > s + f - g(e) \Leftrightarrow f < g(e) - p. \quad \text{(P1)}$$

$$A_1 > D_1 (\text{or } A_2 > D_2 \text{ resp.}) \Leftrightarrow s + f - g(e) > -p \Leftrightarrow f > g(e) - (p + s). \quad \text{(P2)}$$

$$D_1 > C_1 (\text{or } D_2 > B_2 \text{resp.}) \Leftrightarrow -p > s - g(e) - p \Leftrightarrow g(e) > s. \quad \text{(P3)}$$

- **Chicken game:** (player 1, $B_1 > A_1 > C_1 > D_1$; player 2, $C_2 > A_2 > B_2 > D_2$)

- $$B_1 > A_1 (\text{or } C_2 > A_2 \text{resp.}) \Leftrightarrow s - p > s + f - g(e) \Leftrightarrow f < g(e) - p. \quad \text{(C1)}$$

- $$A_1 > C_1 (\text{or } A_2 > B_2 \text{resp.}) \Leftrightarrow s + f - g(e) > s - g(e) - p \Leftrightarrow f > -p. \quad \text{(C2)}$$

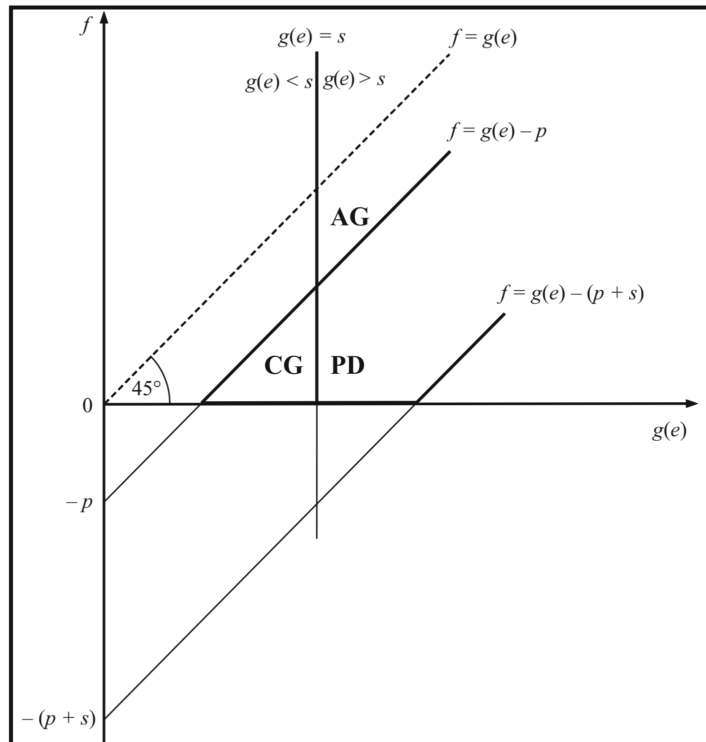
- $$C_1 > D_1 (\text{or } B_2 > D_2 \text{resp.}) \Leftrightarrow s - g(e) - p > -p \Leftrightarrow g(e) < s. \quad \text{(C3)}$$

- **Assurance game:** (player 1, $A_1 > D_1 > C_1$ and $A_1 > B_1$; player 2, $A_2 > D_2 > B_2$ and $A_2 > C_2$)

- $$A_1 > D_1 (\text{or } A_2 > D_2 \text{resp.}) \Leftrightarrow s + f - g(e) > -p \Leftrightarrow f > g(e) - (p + s). \quad \text{(A1)}$$

- $$D_1 > C_1 (\text{or } D_2 > B_2 \text{resp.}) \Leftrightarrow -p > s - g(e) - p \Leftrightarrow g(e) > s. \quad \text{(A2)}$$

- $$A_1 > B_1 (\text{or } A_2 > C_2 \text{resp.}) \Leftrightarrow s + f - g(e) > s - p \Leftrightarrow f > g(e) - p. \quad \text{(A3)}$$



Games, Fig. 5 Efficient areas of the competition game. (Source: Blum et al. 2005, p. 191)

If we plot these results on a graph, we obtain the following result, which is illustrated in Fig. 5. It seems that good competition regimes depend on credible penalties; sufficient externalities for good competitive behavior that can be internalized, i.e., through long-term growth; and a limited disutility of effort, i.e., limited levels of transaction costs for compliance.

What Can Be Learned from Game Theory

Game theory is a formal tool that shows the extent to which the outcomes of different decisions interact, especially if players have opposing goals. It provides a rigorous analytical environment for evaluating efficiency in static or dynamic environments. Although the ability to model reality in a very concrete way is limited, the overall structural messages are extremely important in understanding how to manage dilemma situations. In addition, players should check whether the game is a true abstraction of reality – otherwise, failure can be bitter. In summer 2015, Greek finance minister Varoufakis made this experience when trying in vain to provoke the lender Troika in a poker-type manner and through impulsive behavior in order to improve his negotiation position, a strategy already proposed 50 years ago in the Strategy of Conflict by (Schelling 1960).

Cross-References

▶ Abuse of Dominance
▶ Blackmail
▶ Bounded Rationality
▶ Cartels and Collusion
▶ Coase and Property Rights
▶ Corruption
▶ Credibility
▶ Equilibrium Theory
▶ Governance
▶ Information Deficiencies in Contract Enforcement
▶ Institutional Economics
▶ Rationality

References

Blum U, Dudley L, Leibbrand F, Weiske A (2005) Angewandte Institutionenökonomik: Theorien, Modelle, Evidenz. Gabler, Wiesbaden
Binmore, K., 1992, Fun and Games, D.C. Health Company, Lexington (MA)
von Clausewitz C (1832) Vom Kriege. Dümmlers Verlag, Berlin; (2008) On war. Oxford World's Classics, Oxford
Dixit AI, Nalebuff BJ (2010) The art of strategy, a game theorists guide to success in business and life. Borton, New York
Erhard L (1957) Wohlstand für Alle. Econ, Düsseldorf; (1958) Prosperity through competition. Frederick A. Praeger, New York
Eucken W (1952/1962) Grundsätze der Wirtschaftspolitik. J.C.B. Mohr, Tübingen/Zürich
Hardin G (1968) The tragedy of the commons. Science 162:1243–1248
Huizinga J (1933) Over de grenzen van spel en ernst in de kultur. Rector's speech, Leiden
Kant I (1787) Kritik der reinen Vernunft. Königsberg; (2015) Critique of practical reason, Cambridge texts in the history of philosophy. Cambridge
Nash G (1950) Non-cooperative games, PhD thesis, Mimeo, Princeton
Neumann J (1928) Zur Theorie der Gesellschaftsspiele. Math Ann 100:295–300
Neumann JV, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton
Osborne M, Rubinstein A (1994) A course in game theory. The MIT-Press, Cambridge
Rasmusen E (2006) Games and information: an introduction to game theory. Blackwell, Oxford
Sun Zi (2003) commented by CAO Cao et al. (【春秋】孙武撰,【三国】曹操 等注,宋本十一家注孙子,上海:上海古籍出版社). Shanghai Ancient Books Publishing House, Shanghai
Sun Zi (~500 BC, 2007) The art of war. Columbia University Press, New York
Schelling, Th., 1960 (1981), The Strategy of Conflict, Harvard University Press, Boston (Mass).

GE/Honeywell

Philipp Schumacher
Wildau Institute of Technology, Technical University of Applied Science, Wildau, Brandenburg, Germany

Definition

The GE/Honeywell case, one of the most controversial merger cases in the history of European merger control, was subject to a long and heated transatlantic debate on the vertical and conglomerate aspects of the European Commission's 2001 prohibition decision. However, horizontal overlaps in the specific markets for aircraft and marine engines were, ultimately, the only reason given by the Court of First Instance for upholding the controversial decision.

The GE/Honeywell Case

In 2001, the attempted US$ 45 billion merger between General Electric Company (GE) and Honeywell International, Inc. could have become the largest deal of its kind in industrial history. However, while the U.S. Department of Justice Antitrust Division (DoJ) allowed the merger on the condition that minor structural remedies be carried out, the European Commission prohibited the merger (European Commission 2001a, b). Although the Court of First Instance (CFI) criticized the Commission's analysis of the vertical and conglomerate effects of the merger, it dismissed GE's appeal in 2005, upholding the prohibition decision based on alleged horizontal effects.

GE is a diversified conglomerate and participates in a large variety of markets worldwide. The main industries it served in 2001 were: aircraft engines, transportation systems, industrial systems, power systems, plastics, lighting, medical systems, domestic appliances, broadcasting (via NBC), financial services, and information services. GE's only independent competitors in the supply of civil aircraft engines were Pratt & Whitney, part of United Technologies (USA), and Rolls-Royce (UK). GE also owned one of the world's largest lease aircraft fleets through GE Capital Aviation Services (GECAS).

At time of merger proposal, Honeywell International was a diversified technology and manufacturing company, serving customers worldwide with aerospace products and services; control technologies for buildings, homes, and industry; automotive products turbochargers; and speciality materials. Its aerospace division had a strong position in flight control systems, the

combination of cockpit displays, computers, and software that control the aircraft.

There are a large number of publications available on the GE/Honeywell case which may be categorized by five groups. Firstly, there are official publications from the European Commission, the CFI, and the US DoJ. A second group of literature mainly investigates the divergent outcomes and discusses which merger control system is superior, comprising Varian (2001), Jin (2002), Kolasky (2002a, b, c, d), Burnside (2002), Hochstadt (2003), Emch (2004), Evans and Salinger (2002), Pflanz and Caffara (2002), Ruffner (2003), Morgan and McGuire (2004), Schnell (2004), Fox (2002a, b), Girardet (2006). Thirdly, there is a large body of economic and legal literature focusing on parts of the EC decision. Here, the focus is particularly on *conglomerate effects*, i.e., bundling and leveraging, which received most of the scientific community's attention, (e.g., Briggs and Rosenblatt 2001, 2002; Drauz 2002; Emch 2003; Reynolds and Ordover 2002; Choi 2001; Nalebuff 2002; Grant and Neven 2005; Patterson and Shapiro 2001). The fourth group of literature analyses the ruling of the Court of First Instance comprising Howarth (2006), Montag (2006), Pflanz (2006), Kolasky (2006), Weidenbach and Leupold (2006), Schwaderer (2006, 2007), Fox (2006), von Bonin (2006). Representing the fifth group, Schumacher (2010, 2013) shows that in contrast to the decision of the European Commission and the CFI, the merger would not have led to anticompetitive *horizontal effects*. According to the two-level approach of the European Commission, firstly the competitive situation of the engine manufacturer in relation to its direct customer, the aircraft or marine vessel manufacturer, and secondly the competitive effects in the respective market of end-use applications must be considered. Schumacher (2010) shows that GE was not in the position to exert market power prior to the merger and would not have been ex post. Applying the new EC Merger Regulation of 2004 and its Horizontal and Non-horizontal Merger Guidelines or its 2001 predecessor, an overall competitive appraisal has to go beyond the definition of the relevant market and the calculation of structural indicators such as market shares. Assessment of the competitive closeness, firstly, of the parties' aircraft engines and, secondly, of aircraft powered by these engines requires the analysis of bidding data and technical characteristics. The engine markets affected are worldwide bidding markets characterized by volatile market shares, the high importance of after-sales revenue and profitable outside options. Based on publicly available data, information from aerospace professionals and industry knowledge, the conclusion of an effects-based analysis of GE/Honeywell is diametrically opposed to the Commission's decision regarding both horizontal and nonhorizontal aspects. As shown in similar complex aerospace mergers GE/Smiths Aerospace (2007) and UTC/Goodrich (2012), the "more economic approach" to European merger control, including empirical analysis firmly grounded in industry-specific expertise, can help to avoid a pro-competitive merger being prohibited again based on speculative theories of harm inconsistent with the realities of the case.

Cross-References

- ▶ Efficiency
- ▶ Horizontal Effects
- ▶ Merger Control
- ▶ Merger Remedies

References

Briggs JD, Rosenblatt HT (2001) A bundle of trouble: the aftermath of GE/Honeywell. Antitrust mag 16(1): 26–31

Briggs JD, Rosenblatt HT (2002) GE/Honeywell – live and let die: a response to Kolasky and greenfield. George Mason Law Rev 10(3):459–470

Burnside A (2002) GE, honey, I sunk the merger. Euro Compet Law Rev 23(2):107–110

Choi JP (2001) A theory of mixed bundling applied to the GE/Honeywell merger. Antitrust Mag 16(1):32–33

Court of First Instance (2005) Case T-210/01, General Electric Company vs. Commission of the European Communities. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62001TJ0210&rid=5. Accessed 10 June 2017

Drauz G (2002) Unbundling GE/Honeywell: the assessment of conglomerate mergers under EC

competition law. Fordham Int Law J 25(4): 885–908
Emch E (2003) GECAS and the GE/Honeywell merger: a response to Reynolds and Ordover, Economic Analysis Group Discussion Paper, EAG 03-13. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=451961. Accessed 10 June 2017
Emch E (2004) "Portfolio effects" in merger analysis: differences between EU and U.S. Pract Recomm Future, Antitrust Bull 49(1–2):55–100
European Commission (2001a) Case No. COMP/M.2220 General Electric/Honeywell, OJ (2004) L48/1
European Commission (2001b) The commission prohibits GE's acquisition of Honeywell, Rapid – Press Releases – Europa, July 3. Available from: http://europa.eu/rapid/press-release_IP-01-939_en.htm. Accessed 10 June 2017
European Commission (2007) Case No. COMP/M.4561 GE/Smiths Aerospace, OJ (2007) C133/02
European Commission (2012) Case No. COMP/M.6410 – UTC/Goodrich, OJ (2012) C388/04
Evans DS, Salinger MS (2002) Competition thinking at the European Commission: lessons from the aborted GE/Honeywell merger. George Mason Law Rev 10:489–532
Fox EM (2002a) Mergers in global markets: GE/Honeywell and the future of merger control. Univ Pennsylvania J Int Econ Law 23:457–468
Fox EM (2002b) U.S. and European merger policy – fault lines and bridges. George Mason Law Rev 10:471–488
Fox EM (2006) The European court's judgement in GE/Honeywell – not a poster child for comity or convergence. Antitrust mag 20(2):77–81
Girardet F (2006) Internationales Fusionskontrollrecht – Konflikt und Konvergenz – Eine Untersuchung mit Schwerpunkt auf dem Europäischen und US-amerikanischen Fusionskontrollrecht anhand der Zusammenschlussvorhaben Boeing/MDD und GE/Honeywell. Peter Lang, Frankfurt am Main
Grant J, Neven DJ (2005) The attempted merger between general electric and Honeywell – a case study of transatlantic conflict. J Compet Law and Econ 1(3): 595–633
Hochstadt ES (2003) The Brown Shoe of European Union competition law. Cardozo Law Rev 24:287–395
Howarth D (2006) The court of first instance in GE/-Honeywell. Eur Compet Law Rev 27(9):485–493
Jin P (2002) Turning competition on its head: economic analysis of the EC's decision to bar the GE-Honeywell merger. Northwest J Int Law Bus 23:187–211
Kolasky WJ (2002a) Conglomerate mergers and range effects: it's a long way from Chicago to Brussels. George Mason Law Rev 10:533–550
Kolasky WJ (2002b) United States and European competition policy: are there more differences than we care to admit? Speech at the European Policy Center, Brussels, April 10 Available from: https://www.justice.gov/atr/file/519811/download. Accessed 10 June 2017
Kolasky WJ (2002c) North Atlantic competition policy: converging toward what?, Speech at the BIICL second annual international and comparative law conference, London, May 17. Available from: https://www.justice.gov/atr/file/519801/download. Accessed 10 June 2017
Kolasky WJ (2002d) GE/Honeywell: continuing the transatlantic dialog. Univ Pennsylvania J Int Econ Law 23:513–537
Kolasky WJ (2006) GE/Honeywell: narrowing, but not closing, the gap. Antitrust mag 20(2):69–76
Montag F (2006) Das GE-Honeywell-Urteil des EuG: Absage an konglomerate und vertikale Effekte in der europäischen Fusionskontrolle? In: Brinker I, Scheuing DH, Stockmann K (eds) Recht und Wettbewerb – Festschrift für Rainer Bechtold zum 65. Geburtstag. Verlag C.H.Beck, München, pp 339–355
Morgan EJ, McGuire S (2004) Transatlatic divergence: GE-Honeywell and the EU's merger policy. J Eur Publ Policy 11(1):39–56
Nalebuff B (2002) Bundling and the GE-Honeywell merger. Working paper No. 22, Working paper series ES – Economics/Strategy. Yale School of Management. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=327380. Accessed 10 June 2017
Patterson DE, Shapiro C (2001) Transatlantic divergence in GE/Honeywell: causes and lessons. Antitrust Mag 16(1):18–25
Pflanz M (2006) Economic analysis and judical review: the CFI on GE/Honeywell. Zeitschrift Wettbewerbsrecht 4(2):139–178
Pflanz M, Caffara C (2002) The economics of G.E./Honeywell. Eur Compet Law J 23:115–121
Reynolds RJ, Ordover JA (2002) Archimedean leveraging and the GE/Honeywell transaction. Antitrust Law J 70:171–198
Ruffner T (2003) The failed GE/Honeywell merger: the return of portfolio-effects theory. DePaul Law Rev 52:1285–1231
Schnell DK (2004) All bundled up: bringing the failed GE/Honeywell merger in from the cold. Cornell Int Law J 37:217–262
Schumacher P (2010) Revisiting the General Electric/Honeywell case applying the European Commission's more economic approach. Shaker Verlag, Aachen
Schumacher P (2013) The EU's flawed assessment of horizontal aspects in GE/Honeywell: re-visiting the last pillar of the European prohibition decision. Eur J Law Econ, 2013 35(2):211–240
Schwaderer MA (2006) Anything new after GE/Honeywell. In: Bösche KV, Füller JT, Wolf M (eds) Variationen im Recht – Festbeigabe für Franz Jürgen Säcker. Berliner Wissenschafts-Verlag, Berlin, pp 151–161
Schwaderer MA (2007) Conglomerate merger analysis – the legal context: how the European courts' standard of proof put an end to the ex ante assessment of leveraging. Zeitschrift für Wettbewerbsrecht 5(4):482–514
Varian HR (2001) In Europe, G.E. and Honeywell Ran Afoul of 19th-century thinking, New York Times June 28, C2
Von Bonin A (2006) Vertikale und konglomerate Zusammenschlüsse nach dem Urteil GE/Kommission. Wirtschaft und Wettbewerb 56(5):466–476
Weidenbach G, Leupold H (2006) Das GE/Honeywell-Urteil des EuG – Spannende Rechtsfragen im Überfluss. Europäisches Wirtsch Steuerrecht 17(4):154–163

Gender Diversity

Isabel-María García-Sánchez
Universidad de Salamanca, Salamanca, Spain

From a business perspective, gender diversity represents the heterogeneousness of a workforce in terms of gender, including that of top management and the board of directors. Gender occupational segregation is a phenomenon which is present in most countries, regardless of economic conditions and the existence of anti-discrimination laws that ensure full citizenship for women. However, the presence of women in positions of power and decision-making in organizations and in public and private institutions is virtually nonexistent. Most analysts agree that organizational dynamics are the main obstacle to the career advancement of women and that these are still dominated by androcentric values that exclude them from decision-making. This still does not solve the problem of reconciling work with privacy.

The inclusion criterion for gender diversity in organizations corresponds to a strategy of promoting the participation of women in positions of responsibility. The central idea of the concept of diversity is the maximum utilization of the potential of heterogeneous groups, i.e., how they are different in terms of gender. Variability is emphasized, so that each person is valued for what they are and what they can contribute with regard to their personal characteristics. A diversity strategy gives organizations the ability to attract and retain diverse talents which are representative of both sexes. In this sense, gender diversity recognizes that there is no single way to work but that labor plurality may be advantageous for an organization, as it encourages innovation and complementary perspectives.

Generalized differences were identified in the gender management practices due to women and men having several approaches to overall style, decision-making, and interpersonal relationships. Several authors, such as Robinson and Dechant (1997), suggest that the incorporation of women into the main management bodies entails a significant increase in business due to the economic result set of skills, knowledge, and personal experiences of women. Women have a higher ability to achieve a better understanding of the demands of customers, increasing their chances of entering new markets (Carter et al. 2003), in order to promote an effective resolution of problems and evaluate alternatives (Rose 2007) and to hire the best executives and directors to limit bias against minorities (Campbell and Minguez-Vera 2008). All of this will result in positive signals being sent out to the labor market and the generation of capital for products that will reduce the costs borne by companies (Rose 2007).

Additionally, the proponents of gender diversity on the board feel that their presence is meant to increase business organizational values and business results through the creation of new knowledge and perspectives (Carter et al. 2003). These are the key factors in the process of adopting sustainable forms of behavior in the economic sphere. Furthermore, according to Calas and Smircich (2004), with regard to the traditional feminine qualities, rhetoric is a rationale that must include the value being held for radical changes and for configuring a new business model that creates value not only for shareholders but also for different stakeholders.

The above characteristics are specified in the globally proven fact that women are more sensitive to other perspectives of corporate behavior, have a greater ability to identify the needs of different stakeholders, and can provide more opportunities to satisfy them (Brennan and McCafferty 1997). Thus, counselors need to be more participative and democratic and show more group spirit than their colleagues in promoting participation and dialogue within the board and, by extension, in encouraging greater communication and paying more attention to the interests of the various stakeholders (Eagly et al. 2003). They also need to have had a greater number of experiences outside the business world (Hillman et al. 2002). All of the above creates, ultimately, an increased focus on CSR (Singh and Vinnicombe 2004).

However, from the point of view of value creation, the presence of more women may also involve a number of disadvantages associated

mainly with the diversity of opinions and the appearance of discrepancies that may cause delays in decision-making (Smith et al. 2006).

References

Brennan N, McCafferty J (1997) Corporate governance practices in Irish companies. Ir J Manag 18:116–135
Calas MB, Smircich L (2004) Revisiting 'dangerous liaisons' or does the 'feminine-in-management' still meet 'globalization'? In: Frost PJ, Nord WR, Krefting LA (eds) Managerial and organizational reality: stories of life and work. Prentice- Hall, Upper Saddle River, pp 467–481
Campbell K, Mínguez-Vera A (2008) Gender diversity in the boardroom and firm financial performance. J Bus Ethics 83:435–451
Carter DA, Simkins BJ, Simpson WG (2003) Corporate governance, board diversity and firm value. Financ Rev 38:33–53
Eagly AH, Johannesen-Schmidt MC, van Engen M (2003) Transformational, transactional and laissez-faire leadership styles: a meta-analysis comparing women and men. Psychol Bull 129:569–591
Hillman AJ, Cannella AA Jr, Harris IC (2002) Women and racial minorities in the boardroom: how do directors differ? J Manag 28(6):747–763
Robinson G, Dechant K (1997) Building a business case for diversity. Acad Manag Exec 11(3):21–31
Rose C (2007) Does female board representation influence firm performance? The Danish evidence. Corp Govern An Int Rev 15(2):404–413
Singh V, Vinnicombe S (2004) Why so few women directors in top U.K. Boardrooms? Evidence and theoretical explanations. Corp Govern 12(4):479–488
Smith N, Smith V, Verner M (2006) Do women in top management affect firm performance? A panel study of 2,500 Danish firms. Int J Perform Manag 55(7):569–593

Genocide

Charles H. Anderton[1] and Jurgen Brauer[2]
[1]College of the Holy Cross, Worcester, MA, USA
[2]Hull College of Business, Georgia Regents University, Augusta, GA, USA

Abstract
The entry defines crimes of mass atrocity, such as genocide, provides data on their intensity, and discusses domestic and international institutions formed to address the crimes. In addition, the entry briefly surveys the economic literature on mass atrocity crimes from theoretical and empirical perspectives. It pays particular attention to the economics of international law.

Definition

By combining the Greek *genos* (a people, tribe, race) and the Latin *cide* (to kill), Raphael Lemkin (1944, p. 79) invented the word genocide. Article 2 of the 1948 United Nations (UN) Convention on the Prevention and Punishment of the Crime of Genocide defines *genocide* as "any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, as such: (a) Killing members of the group; (b) Causing serious bodily or mental harm to members of the group; (c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part; (d) Imposing measures intended to prevent births within the group; (e) Forcibly transferring children of the group to another group" (United Nations 1948). As of 31 October 2014, only 146 UN members are Party to the Convention.

Data, Mass Atrocity Crimes, and Institutions

Data

In a summary of large-sample datasets on atrocities involving civilians, Anderton (in Anderton and Brauer forthcoming; henceforth "in A/B forthcoming") identifies 201 distinct cases of state-sponsored genocides and mass atrocities (GMAs) from 1900 to 2013, 43 state-perpetrated genocides from 1955 to 2013, and 34 GMAs perpetrated by non-state groups from 1989 to 2013. Some well-known genocides include the Armenian genocide (1915–1918; estimated fatalities ~1.5 million), the Holocaust (1933–1945; ~10 million), Cambodia (1975–1979; ~1.9 million), Rwanda (1994; ~0.8 million), and Sudan-Darfur (2003–2011; ~0.4 million). A cautious estimate of

intentional civilian fatalities associated with the 202 state-perpetrated GMAs since 1900 is 84 million. Less well-known are non-state perpetrated atrocities such as conducted by the so-called Islamic State, with estimated fatalities of 8,198 from 2005 to 2013 (see Uppsala Conflict Data Program at http://www.pcr.uu.se).

Mass Atrocity Crimes

As defined in the UN Convention, *genocide* is the intentional destruction, in whole or in part, of a specific group of people. In non-genocidal *mass killing*, perpetrators do not seek to destroy a group as such (Waller 2007, p. 14). *Crimes against humanity* encompass widespread or systematic attacks against civilians involving inhumane means such as extermination, forcible population transfer, torture, rape, and disappearances. *War crimes* are grave breaches of the Geneva Conventions including willful killing, torture, willfully causing great suffering or serious injury, and extensive destruction and appropriation of property. *Ethnic cleansing* is the removal of people of a particular group from a state or region using means such as forced migration or mass killing (Pergorier 2013). *Violence against civilians* (VAC) can incorporate mass atrocities but it also includes incidents that are relatively small, specifically, less than 1,000 per case or per year. Along with genocide, crimes against humanity, war crimes, and ethnic cleansing are both legal and scholarly terms.

At the Nuremberg trials of 1945–1946, the International Military Tribunal found none of the accused guilty of crimes committed prior to the outbreak of war on 1 September 1939; litigation was limited to atrocities during wartime (Schabas 2010, pp. 126–127). Lemkin's (1944) and the UN's conceptions of genocide were novel precisely because they spoke to criminal acts committed in wartime or in peacetime. Nevertheless, the UN definition of genocide has been subject to critical scrutiny by scholars, for instance in regard to groups left out (e.g., political), how to identify "intent," the inability of the Convention to prevent genocide, the relationship of genocide to other atrocities, and misuse of the term (e.g., Curthoys and Docker 2008). The UN definition of genocide has not expanded since 1948 to include other groups, but international criminal law has evolved. Schabas (2010, p. 141) maintains that the expanded concept of crimes against humanity has "emerged as the best legal tool to address atrocities" and "genocide as a legal concept remains essentially reserved for the clearest cases of physical destruction of national, ethnic, racial, or religious groups."

International and Domestic Institutions

The twentieth and twenty-first centuries display the emergence and growth of international and domestic laws designed to prevent, punish, and/or foster restitution for atrocity crimes. Table 1 shows a selection of such institutions as well as sources that provide further information. Adjudication of mass atrocity crimes began in earnest following World War I with the establishment of the Turkish Military Tribunal (TMT) (1919–1920), which prosecuted organizers of the Armenian genocide. The trials, characterized as "a milestone in Turkish legal history" (Dadrian 1997, p. 30), revealed the systematic planning behind the genocide, enrichment of perpetrators through looting of victims' assets, and the lack of military necessity for the forced relocation of Armenians. However, the TMT convicted only 15 men among the hundreds who orchestrated the genocide (Dadrian 1997).

Following World War II, the International Military Tribunal (IMT) at Nuremberg was established in which leading officials were tried for war crimes and crimes against humanity (1945–1946). Twelve Nazi leaders received the death sentence and many others were given long jail terms. The trials had an important influence on the growth of international criminal law including the 1948 Genocide Convention, the International Criminal Tribunal for the Former Yugoslavia (ICTY) established in 1993, the International Criminal Tribunal for Rwanda (ICTR) established in 1994, and the International Criminal Court (ICC) ratified in 2002. As of December 2014, the ICTY had indicted 161 people for atrocity crimes associated with the wars in the former Yugoslavia in the 1990s. As of December 2014, the ICTR had indicted 95 people for atrocity

Genocide, Table 1 Selection of legal institutions, jurisprudence, and international norms related to genocide prevention and post-genocide justice

Selection of legal institutions (or norms)	Selection of sources for further information
International	
International Military Tribunal at Nuremberg (IMT) (1945–1946)	US Holocaust Memorial Museum (http://www.ushmm.org)
Convention on the prevention and punishment of the crime of genocide (1948, 1951)	United Nations (https://treaties.un.org), Schabas (2010), US Holocaust Memorial Museum (http://www.ushmm.org)
International Criminal Tribunal for the former Yugoslavia (ICTY) (1993)	United Nations (http://www.icty.org)
International Criminal Tribunal for Rwanda (ICTR) (1994)	United Nations (http://www.unictr.org)
International Criminal Court (ICC) (2002)	International Criminal Court (http://www.icc-cpi.int)
Norms on the responsibilities of transnational corporations and other business enterprises with regard to human rights (2003)	Hillemanns (2003)
Extraordinary Chambers in the Courts of Cambodia (ECCC) (2003)	Extraordinary Chambers in the Courts of Cambodia (http://www.eccc.gov.kh/en)
Responsibility to Protect (R2P) (2005)	United Nations (UN A/RES 60/1 www.un.org/Docs; http://www.un.org/en/preventgenocide)
Domestic	
Turkish military tribunal (1919–1920)	Dadrian (1997)
US Alien Tort Claims Act (1789, 1980)	Michalowski (2013)
Prosecution of civilian atrocities (not necessarily genocide) in domestic courts (includes Nuremberg and others)	Schabas (2003), Prevent genocide international (http://www.preventgenocide.org)

crimes associated with the 1994 civil war and genocide and it established the legal precedent that mass rape during wartime is genocidal. Following the huge backlog of cases awaiting trial in Rwanda, the government turned to the Gacaca court system, based on traditional law developed within communities (Bornkamm 2012). As of October 2014, the ICC has indicted 36 individuals for atrocity crimes including three current or former heads of state: Omar al-Bashir (Sudan), Uhuru Kenyatta (Kenya), and Laurent Gbagbo (Côte d'Ivoire). According to its Statutes, the ICC has jurisdiction with respect to genocide, crimes against humanity, and war crimes.

Another important international genocide development occurred at the 2005 UN World Summit, in which member states unanimously adopted a norm known as the *Responsibility to Protect* (R2P). R2P was part of the impetus for UN Security Council Resolution 1973, passed on 17 March 2011, which authorized member states to take actions, including enforcement of a no-fly zone, to protect civilians from attacks by the Libyan military. Nevertheless, the UN's R2P resolution has no legal force (UN Doc. A/RES/60/1, paras 138, 139).

Following the Nuremberg trials and the UN Convention, several dozen nations have developed domestic laws to put on trial suspected Nazi war criminals and/or perpetrators of more recent atrocities (Schabas 2003). For example, in 2000 the Chilean Court of Appeals lifted former President Augusto Pinochet's immunity from prosecution, paving the way for trial for his role in civilian atrocities that occurred during his leadership (Pinochet died prior to any conviction). The case is notable not only because it involved a state's prosecution of its former leader, but also because Pinochet's initial arrest occurred in London based on an application of "universal jurisdiction" by European judges. Universal jurisdiction is a principle by which a state (or states, in the Pinochet case) asserts its right to prosecute a person for an alleged crime regardless of the crime's location and the accused's residence or nationality (Lunga 1992).

Not shown in Table 1 are formalized norms within for-profit and non-profit organizations designed to inhibit complicity in atrocities. The ICC followed the ICTY and ICTR in having jurisdiction only over "natural persons" and not "legal persons" (Cernic 2010, p. 141), which ruled out prosecution of corporations complicit in genocide (individual agents within corporations can be tried). Multinational corporations have been complicit in genocide in many cases, but have not usually faced prosecution (Kelly 2012). Nevertheless, there have been legal efforts, including use of the US Alien Tort Claims Act, to bring litigation against corporations for alleged complicity in atrocities and other human rights abuses. Such litigation is leading companies to develop norms to avoid such complicity (Michalowski 2013).

Theoretical and Empirical Aspects

This section asks: What are (some of) the risk factors that economic theory and associated empirical work point to, i.e., what makes the risk nonzero ($r > 0$)? The section thereafter asks: How can genocide risk be reduced below 1 ($r < 1$)?

Theoretical Perspectives

Formal economic models of genocide are relatively new in the literature on conflict, peace, and security between and within states. Verwimp (2003), Ferrero (2013), Anderton and Carter (2015), and Anderton and Brauer (in A/B forthcoming) present nonstrategic constrained optimization models to highlight conditions under which a political authority would choose genocide as part of its goal of controlling territory or government (or both). The models reveal conditions in which genocide has a low opportunity cost for the authority, specifically, when genocide enhances the authority's control in the context of crisis or war, is not too disruptive to economic activities (e.g., trade), is conducive to looting victims' wealth, and is not likely to generate third-party intervention. Under such conditions, genocide is "cheap," and a positive amount demanded can exist. Genocide prevention requires that the opportunity cost of genocide is made high through sanctions, credible threats of third-party intervention to help victims and/or oppose authorities, threats of prosecution, and surveillance of atrocities which can lead to "naming and shaming" of perpetrators. In addition to modeling genocide risk factors, Anderton and Brauer (in A/B forthcoming) use a Lancaster household production model to study the "optimal" choice of genocidal techniques (e.g., mass killing, starvation, forced relocation, etc.) by a regime that has already chosen genocide. Among the results is a "bleakness theorem" in which protection policies along one or just a few dimensions have relatively little overall effect, and sometimes no effect, in protecting victims.

Game theory models of genocide consider strategic interactions between warring groups and/or between an oppressive in-group and an out-group in which intentional destruction of civilian groups is part of war tactics or strategy. For example, Azam and Hoeffler (2002) identify conditions in which warring sides use violence against civilians to strengthen themselves in their strategic interaction. Focusing on the years preceding the 1994 Rwandan genocide, Verwimp (2004) develops a four-player game to model the strategic interactions among the regime, the domestic opposition, a violent rebel group, and the international community. Within the game, eliminating the moderate Hutu opposition and exterminating the Tutsi can be "optimal" strategies. Anderton (2010) draws upon the bargaining theory of war to show how severe threat against an authority group or an incentive to eliminate a persistent rival can lead to genocide as an "optimal" choice. Anderton (2010) and Gangopadhyay (in A/B forthcoming) use evolutionary game theory to model how genocide can become socially contagious (acceptable) among "ordinary people." Vargas' (in A/B forthcoming) model of contestation between a government and a rebel group reveals the incentives of each to side to kill the civilians who are supporting the enemy. Within the model, Vargas finds that the strengthening of either side can have ambiguous effects on the total number of civilians killed, thus showing that third-party support for one side or the other can potentially increase civilian killing. Esteban et al.'s (forthcoming; in A/B forthcoming) inter-temporal models of

contestation between a government and a rebel group reveal several important and sometimes counterintuitive results. In particular they find that new discoveries of resources, democratization of the polity, and third-party intervention to defend vulnerable civilians can enhance incentives for mass killing if they materialize under the “wrong” conditions.

The lessons of constrained rational choice and game theory models for thinking about the emergence of laws designed to punish and prevent genocide are critical. Laws that come into being will be evaluated by potential perpetrators of genocide as part of the constraint set being faced. Such agents, if determined to carry out genocide, have an extensive menu of inputs for working around such laws to achieve objectives. Laws to punish or prevent genocide must consider the multiple options available to potential perpetrators and the potential for laws to lead to unintended consequences. This concern is especially significant in the context of strategic interplay between a government, rebel organization, and possible third-party intervener. If not carefully designed, law to prevent and punish genocide can serve to increase incentives for atrocities. (On the design of law, see the next section.)

Perspectives from behavioral economics also help to study genocide (e.g., Anderton and Brauer; Slovic, Västfjäll, and Gregory, both in A/B forthcoming). Especially important is the reference-dependent objective function of one or a few leaders who have become accustomed to control of political, economic, and/or territorial goods. Experiments in behavioral economics often find evidence of *loss aversion*, in which, relative to a reference point such as current hold on power, subjects believe they are worse off from a loss than a similar gain leads them to feel better off. The notion of loss of power as a form of extreme crisis or existential threat in the minds of leaders is palpable in many genocide case studies (Totten and Parsons 2013). Such losses, coupled with the behavioral phenomenon of loss aversion, suggest that leaders could make extreme choices including repressive violence or genocide to avoid loss (Midlarksy 2005, pp. 64–74).

Empirical Perspectives

About 30 published large-sample cross-country empirical studies of genocide or other forms of VAC risk or seriousness exist (see Anderton 2014; Anderton and Carter 2015; and Hoeffler in A/B forthcoming). Most of these focus on genocide risk or severity from the perspective of *countries*, and thus they focus on the problem of genocide from the “macro” or top-down perspective. Another branch of empirical genocide literature focuses on particular countries, regions, or locales in which genocide took hold and spread, thus emphasizing a “micro” or bottom-up perspective. While almost all of the empirical studies of genocide in the literature focus on risk or seriousness based on historical data, studies are emerging with an emphasis on forecasting (e.g., Rost 2013; Butcher and Goldsmith in A/B forthcoming).

The most prominent macro-empirical study of genocide risk in the literature is by Harff (2003), who focused on a sample of states that experienced “state failure” (e.g., civil war, regime collapse) from 1955 to 1997. Of 126 state failures in the sample, 35 led to genocide. Conditioned on state failure, Harff used logit analysis to identify six significant risk factors for genocide onset: magnitude of political upheaval; history of prior genocide; exclusionary ideology held by the ruling elite; autocratic regime; ethnic minority elite; and low trade openness. Failing to make the list of significant risk factors was economic development, which Harff proxied by infant mortality. Another important macro-empirical study of mass atrocity risk is Easterly et al. (2006), who assemble a dataset for many countries for the period 1820–1998. Among their key results, they find that mass atrocity is significantly less likely at high levels of democracy and economic development, in which the latter was proxied by real income per capita.

Regarding potential *economic* risk factors for genocide, subsequent empirical research suggests that Harff’s result for trade is not robust with most studies reporting no significant impact of trade on atrocity risk. In addition, Harff’s result on economic development is open to question because an inverse relationship between real income per capita and atrocity risk or seriousness is one of the few modest empirical regularities in the literature.

Other economic risk factors considered in the empirical literature are income inequality and resource dependence, in which no empirical regularities have yet emerged, and *economic* discrimination, which is only beginning to be considered but in which two studies report a significant positive effect on genocide risk (Rost 2013; Anderton and Carter 2015). In the emerging empirical forecasting literature on genocide, no clear results as yet stand out in regard to the roles of economic variables. In their survey of such literature, Butcher and Goldsmith (in A/B forthcoming) suggest that "while economic factors might have an underlying causal effect on the likelihood of genocidal violence in a society, as predictors in forecasting models they might be overshadowed by political or demographic factors that are more proximate to genocide onset."

In addition to large-sample, macro-empirical studies of genocide risk or severity are country-specific, micro-empirical studies that focus on particular characteristics of a nation, region, or individuals that led to the onset or spread of atrocity (e.g., Ibañez and Moya in A/B forthcoming; Justino in A/B forthcoming). Country-specific studies typically identify historical, social, and economic conditions particular to the country and tactical and strategic aspects of war that are critical for understanding atrocity. Such finer-grained elements can be glossed over in large-sample cross-section studies. For example, Ibañez and Moya's (in A/B forthcoming) study of Colombia reveals many dynamic, nuanced, and interrelated aspects of community and household incentives for civilians to flee violence, why some do not flee, and why contesting military forces (government, militias, rebels) tactically and strategically kill and/or force the relocation of civilians. Such complexities are obviously critical in considering laws to reduce risks of future civilian atrocities, but also in prosecuting perpetrators and designing reparations in post-genocide settings.

Economics of International Law

Despite some cases of GMA having been brought to trial in national and international courts or tribunals – the Armenian trials in Turkey, the Nuremburg trials, the Pinochet case, and more recent tribunals regarding Cambodia, Rwanda, and the Balkan wars of the 1990s – the overall record of reducing the risk of GMA to below certainty ($r < 1$) is only mildly encouraging. There are several reasons for this. First, even assuming away issues of ignorance and apathy, as a matter of *economics*, unilateral action runs into the problem of sufficient scale and multilateral, collective action into issues related to strategic behavior, free-riding, coordination, agency, benefit appropriation, and cost shifting. Even assuming that none of these pose a problem, all options rely on the existence of well-codified and well-functioning regimes of national and international law and their enforcement. Second, as a matter of *law*, then, reducing GMA risk is difficult because (a) state sovereigns are cautious to accede to any international treaty that may later expose them to legal liability in the first place and (b) because state sovereigns generally do not cede jurisdiction over nonstate GMA actors to international bodies (e.g., Nigeria maintains jurisdictional prerogative over Boko Haram; and if a nonstate actor prevails in an internal conflict it may not be brought to justice at all). And third, as a matter of *institutional design*, these topics bring up questions, to echo Oliver Williamson (1999), as to what kind of bad GMAs are in the first place and, correspondingly, what kind of good GMA-related laws are, and how to best supply them.

On the demand (or usage) side, are GMA and GMA-related law private (excludable and rivalrous), public (nonexcludable, nonrivalrous), club (excludable, nonrivalrous), or common-resource pool (nonexcludable, rivalrous) bads or goods, or some changing mixture thereof? And on the supply side, are they best provided by private or public actors, or some changing combination of the two, and what is the technology of their production (e.g., best-shot, weakest-link, aggregate effort, or variants thereof)? What sort of issues in agency, transaction costs, and institutional design arise? While a considerable global public goods (GPG) literature has sprung up in economics (e.g., Kaul and Conceição 2006 and literature cited

therein), application to the design of international law as an instance of GPGs is thin in general and almost entirely absent in regard to law and GMA (see, e.g., a recent symposium of papers in the European Journal of International Law, 23(3), 2012).

As regards GMAs, we suggest that indiscriminate chemical weapons gasing may be conceptualized as a *public bad* for the affected population if it is neither feasible to exclude oneself from the gassing nor feasible to seek effective shelter (there can be no rivalry for shelter if there is none to be had). Those who do manage to crowd into a shelter, however, partake in the benefit it offers, the shelter being a common-resource pool good, while those left behind on the street suffer a *common-resource pool bad* (nonexclusionary but rivalrous). In contrast, genocide would be a *club bad* precisely because its architects differentiate and select victims. Finally, examples of a *private bad* suffered in violent conflict include un-orchestrated rape in war or the death of a soldier in the performance of his or her duties (the "expected" bad in war, but not a war crime). Similarly, in regard to the good that GMA-related law may provide, international law of war is intended as a GPG in that all soldiers share in the benefits the law provides and none of them are excluded. In contrast, national law is private to the state whose legislative body passes it: It excludes nationals of other states and reserves benefits to its own nationals. But all international law is effectively a club good, benefitting those who accede, and becomes a pure GPG if and only if *all* states become Party to the treaty in question. In practice, however, it is conceivable that benefits can be withheld so that the benefits law offers become rival to those with the means to access its provisions when needed. Thus, while the Genocide Convention is (not quite) a global public good in principle, the evident practice of "too little, too late" suggests that its enforcement is rivalrous and therefore constitutes a common-resource pool good.

Even this cursory "walk around goods space" (Brauer and van Tuyll 2008, Chap. 8) suggests that the good (or bad) in question can take various forms and that each may change across geographic space and time. Neither the goods nor the bads are necessarily unitary (of a single form), and to conceive of GMA simply as a global public bad requiring a global public good response may be inadequate. Moreover, as Shaffer (2012) points out, international laws can be rivalrous to each other and their construction is designed, in part, to trade off against multiple national laws (legal pluralism).

In addition, economically efficient (no under- or overprovision) GMA-law in response to GMAs may depend on the summation technology of GMA production. Applying Hirshleifer's (1983) insight, that some GPGs are best provided as best-shot products (the single-best effort suffices; no need for anyone else to contribute to its provision), weakest-link products (the weakest provider limits the good's effectiveness), or aggregate effort products (the more is provided by all, the better for all), Shaffer (2012; esp. Table 2, p. 690), argues that best-shot GPGs are best dealt with in global administrative law, weakest-link GPGs by fostering legal pluralism, and that only aggregate effort GPGs may require a global constitutionalist approach. To illustrate, when a single country has effectively become the world's only superpower to intervene in other states' (GMA or GMA-alleged) affairs, it may be tempted to overreach or under-reach according to its own cost-benefit calculated perception of its Responsibility to Protect, regardless of the wishes of all other UN members. But superpower intervention or nonintervention solely at its own discretion challenges global legitimacy (the US is often accused in this regard; France, in regional interventions, less so). Such situations, Shaffer (2012) argues, are best dealt with by global administrative law which might hold the "incumbent" of the superpower "office" responsible for its actions. We imagine (since Shaffer does not address GMA), that instead of a Genocide Convention, there might exist an UN-approved automatic trigger obligating the superpower to intervene in cases of GMA, subject to global administrative law. As of this writing, little has been theorized in this regard.

An additional issue pertains to transgenerational global public goods. Again, this is

insufficiently theorized but probably of great importance in cases of GMAs since each event carries significant generational implications (for a review see, e.g., Ibañez and Moya in A/B forthcoming). For public goods provision, Sandler (1999) speaks for four levels of awareness rules: First, the myopic view considers making a marginal cost (MC) contribution to the provision of a GPG only up to the sum of the marginal benefits (MB) a state estimates for its own current generation, $MC = \Sigma MB$. Second, although still selfish, a forward-looking view is to include one's own offspring generations, i, such that $MC = \Sigma MB_i$. Since the expected benefits are larger, this translates into greater willingness to make a larger MC contribution. Third, a more generous view of the benefits summation includes other states' populations, j, but only for the current generation ($MC = \Sigma MB_j$). The most enlightened view of all – we call this the "Buddha rule" – sums the expected benefits across all generations across all populations, $MC = \Sigma MB_{ij}$. Since such benefit is likely to be large, it justifies correspondingly large outlays.

Finally, design criteria for GPG that would take account of goods (or bads)-space, summation technologies, transboundary, and transgenerational aspects have been discussed in the literature (Sandler 1997) but rarely in regard to GMA-related national and international law (Myerson, in A/B forthcoming, is an exception). It would appear that a fruitful field of inquiry is ready for exploration in this regard.

References

Anderton CH (2010) Choosing genocide: economic perspectives on the disturbing rationality of race murder. Def Peace Econ 21(5–6):459–486

Anderton CH (2014) A research agenda for the economic study of genocide: signposts from the field of conflict economics. J Genocide Res 16(1):113–138

Anderton CH, Brauer J (eds) (forthcoming) Economic aspects of genocide, mass atrocities, and their prevention. Oxford University Press, New York

Anderton CH, Carter J (2015) A new look at weak state conditions and genocide risk. Peace Econ Peace Sci Public Policy 21(1):1–36

Azam J, Hoeffler A (2002) Violence against civilians in civil wars: looting or terror? J Peace Res 39(4):461–485

Bornkamm PC (2012) Rwanda's Gacaca courts: between retribution and reparation. Oxford University Press, New York

Brauer J, van Tuyll H (2008) Castles, battles, and bombs. The University of Chicago Press, Chicago

Cernic JL (2010) Human rights law and business: corporate responsibility for fundamental human rights. Europa Law Publishing, Groningen

Curthoys A, Docker J (2008) Defining genocide. In: Stone C (ed) The historiography of genocide. Palgrave Macmillan, New York, pp 9–41

Dadrian VN (1997) The Turkish Military Tribunal's prosecution of the authors of the Armenian genocide: four major court-martial series. Holocaust Genocide Stud 11(1):28–59

Easterly W, Gatti R, Kurlat S (2006) Development, democracy, and mass killing. J Econ Growth 11(2):129–156

Esteban J, Morelli M, Rohner D (forthcoming) Strategic mass killings. J Polit Econ

Ferrero M (2013) You shall not overkill: substitution between means of group removal. Peace Econ Peace Sci Public Policy 19(3):333–342

Harff B (2003) No lessons learned from the Holocaust? Assessing risks of genocide and political mass murder since 1955. Am Polit Sci Rev 97(1):57–73

Hillemanns CF (2003) UN norms on the responsibilities of transnational corporations and other business enterprises with regard to human rights. Ger Law J 4(10):1065–1080

Hirshleifer J (1983) From weakest-link to best-shot: the voluntary provision of public goods. Public Choice 41(3):371–386

Kaul I, Conceição P (eds) (2006) The new public finance: responding to global challenges. Oxford University Press, New York

Kelly MJ (2012) Prosecuting corporations for genocide under international law. Harvard Law Policy Rev 6(2):339–367

Lemkin R (1944) Axis rule in occupied Europe: laws of occupation, analysis of government, proposals for redress. Carnegie Endowment, Washington, DC

Lunga L (1992) Individual responsibility in international law for serious human rights violations. Springer, New York

Michalowski S (ed) (2013) Corporate accountability in the context of transitional justice. Routledge, New York

Midlarksy MI (2005) The killing trap: genocide in the twentieth century. Cambridge University Press, New York

Pergorier C (2013) Ethnic cleansing: a legal qualification. Routledge, New York

Rost N (2013) Will it happen again? On the possibility of forecasting the risk of genocide. J Genocide Res 15(1):41–67

Sandler T (1997) Global challenges. Cambridge University Press, New York

Sandler T (1999) Intergenerational public goods: strategies, efficiency and institutions. In: Kaul I, Grunberg I, Stern MA (eds) Global public goods:

international cooperation in the 21st century. Oxford University Press, New York, pp 20–50
Schabas WA (2003) National courts finally begin to prosecute genocide, the 'crime of crimes'. J Int Crim Justice 1(1):39–63
Schabas WA (2010) The law and genocide. In: Bloxham D, Moses AD (eds) The Oxford handbook of genocide studies. Oxford University Press, New York, pp 123–141
Shaffer G (2012) International law and global public goods in a legal pluralist world. Eur J Int Law 23(3):669–693
Totten S, Parsons WS (eds) (2013) Centuries of genocide: essays and eyewitness accounts, 4th edn. Routledge, New York
United Nations (1948) Convention on the prevention and punishment of the crime of genocide. https://treaties.un.org
Verwimp P (2003) The political economy of coffee, dictatorship, and genocide. Eur J Polit Econ 19(2):161–181
Verwimp P (2004) Games in multiple arenas and institutional design on the eve of the Rwandan genocide. Peace Econ Peace Sci Public Policy 10(1):1–47
Waller J (2007) Becoming evil. Oxford University Press, New York
Williamson O (1999) Public and private bureaucracies: a transaction cost economics perspective. J Law Econ Organ 15(1):306–342

Geodesy

Otmar Schuster
Haus der Geoinformation, Mülheim an der Ruhr, Germany

Abstract

Geodesy, starting as a scientific discipline exploring the size of the earth in the eighteenth century, soon produced the basis of modern land tax systems and enabled the institution of individual property rights in real estate, a backbone of the constitutional order. Geodesy stimulated the development of the modern state by data collection and land management methods. The satellite techniques and the advancement and the expansion of the sensor technology based upon the theoretical mastery of the space in methods and incoming data made the geodesy and its professionals effective in the public management and the economy. Geodesy as an economic sector producing and managing geo-information is deeply woven in different disciplines of science, public administration, and economy but recognizable by its methods and results.

Definition

Geodesy is a discipline in science, government, land management, and industry. It addresses the shape and the shape near earth, the national government of land, the land management, and real estate property.

Geodesy: Reference Frame for the Earth

Geodesy explains the knowledge about the shape and shape near earth. It comprises specific methods which are useful to analyze tectonic movements but also to manage countries, land, and plots.

It needed more than two millennia to go back to the anthropocentric worldview, but then in the seventeenth and more in the eighteenth century, the technical progress blessed Europe with telescope and a growth of mathematics. The conquest of the continents brought new knowledge about the earth. The question for the right dimension and the shape of the earth moved into the center of the scientific research. Since the angular measurement succeeded to feature a considerable accuracy, the measurement of length was the biggest problem for 200 years (Torge 2009).

The geodetic science succeeded to improve the determination of the dimension of the earth by long triangulation chains along the several meridians on the northern and southern hemisphere. The determination of the International Meter was more or less a by-product. On the basis of this and the expansive surveying methods, the European governments had new tools to manage the countries in a just way. The national economies were able to flower out in an unimagined way.

This deconvolution was a precondition for the preeminence of the European nations in the world. The implemented techniques gave distinction to the scientific specialization under the name of

G

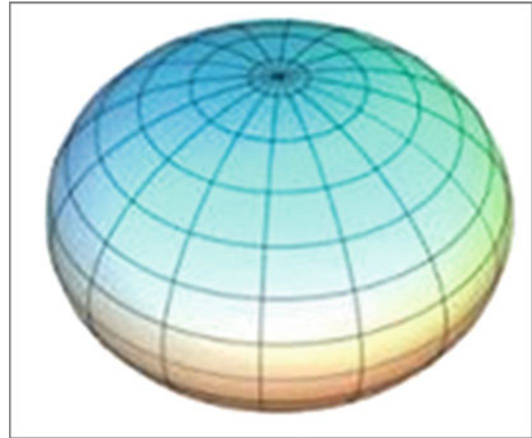
"geodesy." The application of the methods in small-scale areas of about less than 1 km^2, in which the spatial discrepancies are negligible to small for a sufficient accuracy (Kahmen 2005), coined the profession of the "geometer," later referred to as "surveying engineer".

Whereas the systems of base mapping – big or small – got a national character, because the military security asked for that, geodetic technology always had to be international. Only with international cooperation and in addition with the cooperation with the neighbored geosciences, the possibility evolved for tackling the phenomenon "earth" (Ledersteger 1969) (Fig. 1). And all the more so when the scientific progress bestowed methods on us which could determine the gravity field, the rotation of the earth, the movements of the continents, and the mass movements of the earth's interior (Fig. 2).

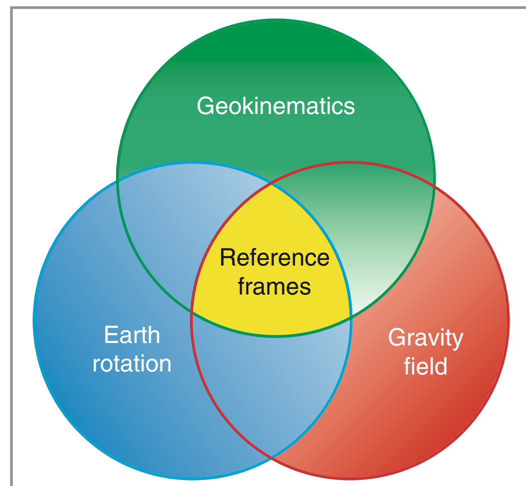
To determine the geodetic frame for national control networks was always one of the most important tasks of geodesy. But international cooperation on this became more and more successful (Schuster 2005). The International Earth Rotation and Reference Systems Service (IERS) maintains the International Terrestrial Reference System (ITRS), which describes the procedures for creating reference frames suitable for use with measurements on or near the earth's surface. This International Terrestrial Reference Frame consists of the high-precision coordinates and velocities of about 400 globally distributed stations (Fig. 3). These three-dimensional Cartesian coordinates and velocities define the positions and the plate tectonic movement, the ITRS.

The ITRF points are part of the European ETRF and the regional networks like the German control net, called DREF, or even networks of lower order under national control; they form together the international frame of high precision. It makes us measure the dislocations of the 19 tectonic plates (Fig. 4), (Sella et al. 2011).

From the beginning astronomical geodesy was necessary to locate the datum of every triangulation net on the ellipsoid. Soon satellite orbits were calculated, and the theoretical earth tides were compared with those measured by gravimeters and horizontal pendulums. Insofar geodesy was



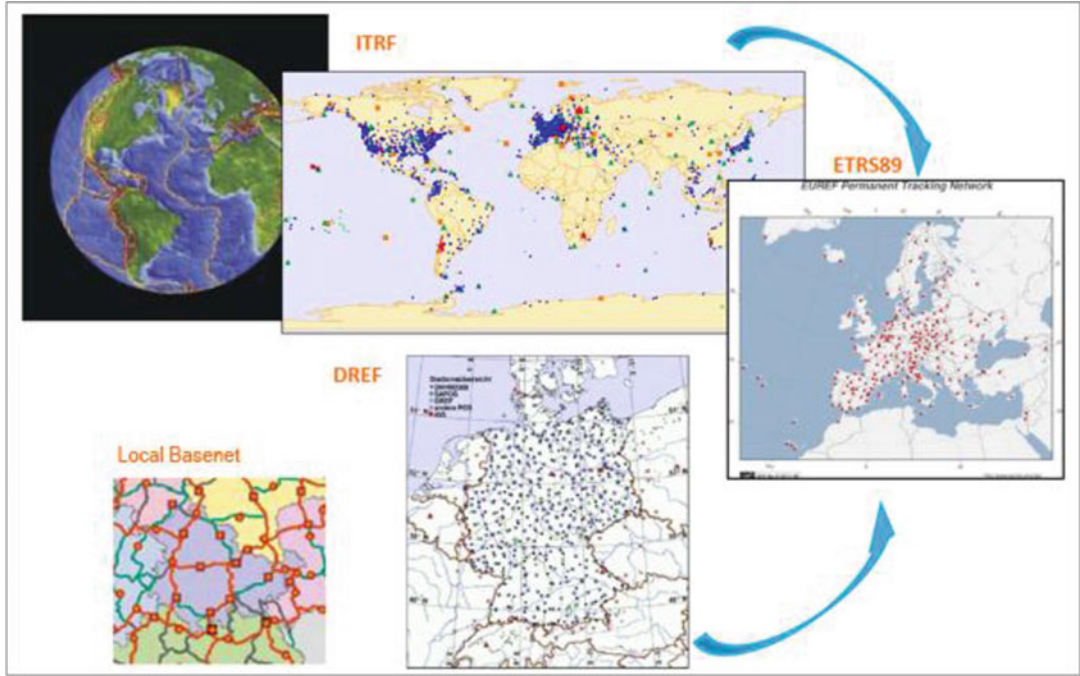
Geodesy, Fig. 1 The earth as a mathematical abstraction



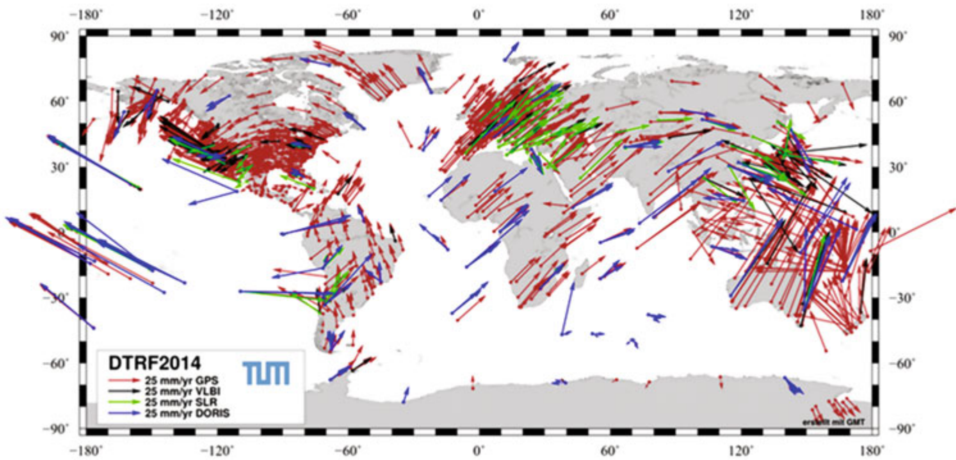
Geodesy, Fig. 2 Study object earth

prepared for the huge next step, the use of the earth near satellites. Up to that time, the meteorological conditions only were disruptive elements for the optical determination of lengths and heights, but then the meteorological surroundings more and more came into view with the higher accuracy of all results.

Geodetic measures are deeply influenced by the shift of masses (Fig. 5) between tectonic plates and the respective interdependence with solid earth and oceans, atmosphere, hydrosphere, cryosphere, biosphere, and asthenosphere (universe). The scientific picture of the earth changed from the spherical shape of the spheroid to the ellipsoid, further to the



Geodesy, Fig. 3 The shape of the earth – under control (Courtesy to Prof. Harald Schuh, GFZ Potsdam)

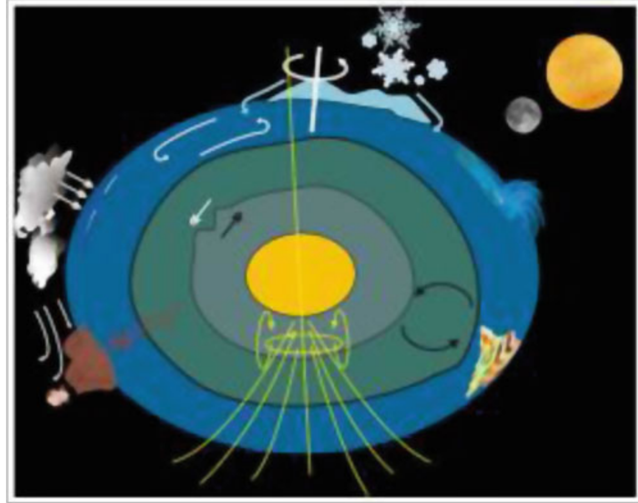


Geodesy, Fig. 4 Tectonic plate dislocations (ITRS Combination Centre at DGFI-TUM 2014) https://www.dgfi.tum.de/international-services/itrs-combination-centre/

geoid. In addition the geodetic science taught us that our earth is in an ongoing process of change.

Today, the highly accurate geodetic frames ITRF and ETRF and the national frames like DREF are combined, even merged. Everybody can locate himself or move in that system. His measurements – mounted in the net – are now worldwide identifiable for about better than 1 [cm] (Seitz et al. 2016). Unlike earlier times the surveying engineer can detect, if the sunspot

Geodesy, Fig. 5 Opposing forcing affecting the earth (Vienna University of Technology)



activity is strong and the magnetic field is jammed. This he has to take care of when performing his measurements.

Individual Property Rights

During the colonization of earth, people became increasingly aware that fruitful land is a scarce resource. It is worth to settle there and to assert the ground. On those places, where people met for trade and change, regulations were accepted for the peaceful way of dealing with each other. Not mentioning the tremendous achievements of the Romans in surveying of their empire, the geodesy steps into the limelight of history, when the harvest tax systematically turned to ground tax. At the end of the eighteenth century, the science provided the technological possibility to expanse scalable maps over the country and to document securely the property borders (Fig. 6). Since that time a fair, area-based tax regime became possible. Tax cadastres became fashionable for modern states. Soon large-scale soil fertility evaluation systems for farmland were set up on these cadastres.

The systematization of the taxation (Fig. 7) and the spread over the countries took technically and politically nearly a century in Europe. But this installation provided the nations with safe funds (Fig. 8). The other side of the coin was that the citizen's rights in their property were strengthened. Geodetic technique was weaved in the deploying civil system of justice. The civil code, the land register ordinance, and the official regulations for the property cadastre came to their heyday at the beginning of the twentieth century. The taxable value is a legal definition for private plots, operational premises, and farming or forestry property. It should be equivalent to the market value, but all over Europe, there is a great variety of tax procedures and level of taxation. In countries like the Netherlands and Sweden, the method of mass evaluation for the ground taxation is discussed. A lot of member states of the European Union are seeking for an easy and secure way for this taxation. The qualitative classification of the nations concerning just ground tax, secure property, and successful mortgaging is an unmistakable sign for the development of land management.

The development did not come to halt (beside the communistic part of the world). Legal and management principles changed with the development of the economy and its regulations. The improved cartography and copying techniques lead to further cadastres showing other attributes of parcels and their buildings (i.e., cadastre on public land charges or pollution suspicion cadastre, geothermal cadastre, ao). They influence the rights in land substantially.

The geodetic professionals working in private or public businesses managed the property

Geodesy, Fig. 6 Modern cadastral map, 1:1000

cadastre and taxation in form of cadastre measurements and cadastre amendments and the evaluation of the ground (Schuster 1981).

In the nineteenth and twentieth century, the rural and urban land readjustment and the management of settlement became an important economic factor driven by geodetic professionals. With the measures of public authorities, the farmers succeeded to raise their crop yields substantially; the drift of the population to the cities was stopped. They succeeded to overcome the big streams of people into the industrial areas.

At least after the Second World War, they managed to integrate 12 million people into the West German population. Geodetic reallotment and evaluation measures were a significant part of the success.

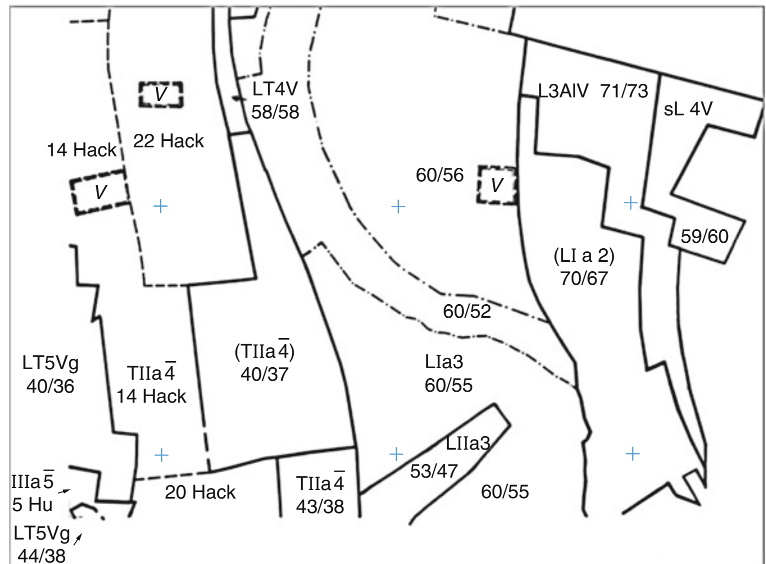
Real Estate Economy

Geodesy is a basis for all real estate economy.

The upcoming time of geo-information opens new economic prospects and effective tools for the protection of our environment. The safe governmental documentation of properties and the loan capabilities (Kerl and Schuster 2003) provided the growing economy by capital generation. The results are reflected in a growing built-up space for industry and residence. Diverse forms of right or property were developed on the basis of the property register and cadastre. They developed their own business cycles. Geodetic services were part of this process of development, but managed mostly only the geometric relocation of the property borders. Slowly they supplied the growing demand of their clients in consulting concerning the more and more complex subject.

In the socialist countries, this development failed even if the geodetic basics and also services for mapping and the building regulations were still or even newly existing. Without the individual property in land, the economic shaping forces were lame. The real estate economy was no area

Geodesy, Fig. 7 Soil fertility valuation, 1:5000



Geodesy, Fig. 8 Soil fertility map, 1:2000

of added value. After the political transformation in 1989, this form of economy on the basis of land quickly induced a takeoff. The authorities quickly transformed their character and got reformed. But with these events, the societal questions around land are not yet answered. The more developed a nation and its civil class is, the more sophisticated and dense is the wood of regulations based on

land. This business life pushes the gross national product of a country immensely. But in that sphere, there is easily coming up the danger of non-genuine growth: economic blows and the growth of bureaucracy can jeopardize the national economy. Both emerge because of inappropriate regulations and misguided public work.

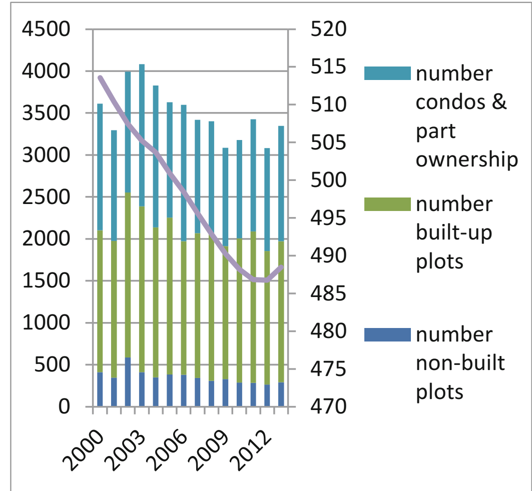
Geodesy is part of this economic process by its land management component, which expanded its methods and performance strongly in the second half of the twentieth century. The German federal building law (BauGB) and the federal land use order (BauNVO) together with the rapidly growing public law considerably influenced the methods of property evaluation and readjustment, which were used in the public and private part of the business life. Other developed countries have similar regulations.

The development of standard ground values and the installation of committees of expert valuers became an internationally acclaimed model for subduing the blows or deflation elements.

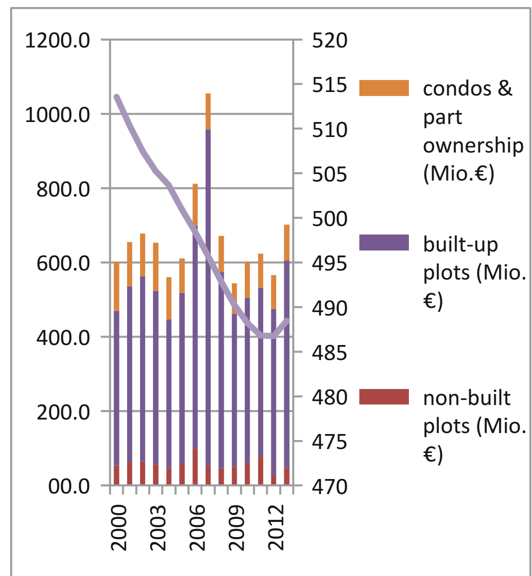
Since the Millennium not only the migration from country to town was noticeable but also the diminishing of an aging European population becomes apparent. The economic output of the house construction industry became much smaller and this happened as well in that part of the geodetic economy. On the other hand, the investments for the maintenance and renewal of the buildings pushed ahead by certain tax reductions and ecological regulations grew strongly. The industry sold their houses for workers by privatization, so the number of transactions swelled (Fig. 9). The people as economic actors wish to safeguard their property, and by that the economic transactions remained on a high level.

Foreign capital searched from 2005 to 2008 for investment opportunities. This pushed the price level especially in the European agglomeration zones. Geodetic services are involved in that part of the economy with evaluation of real estate, property consulting, and evaluation for mortgages.

These different types of propulsions let the transaction traffic remain on a high level, even if the development of new development in unbuilt areas was strongly reduced.



Geodesy, Fig. 9 Numbers of transactions – city of Duisburg



Geodesy, Fig. 10 Cash flows of transactions – city of Duisburg

From both pictures (Figs. 9 and 10), you can deduce the economic importance of the real estate economy. In a city of 500.000 inhabitants happen to be concluded not more than about 4000 contracts per year, which means not more than about 8.000 contract parties. It is a small part of the town's people to exchange a fortune of 500 Mio € up to 1 billion € per year. A number of 250–500

unbuilt plots with a transaction value of 50 to 80 Mio € changed the ownership. These plots are predominantly the basis of a development of the city further on.

The number of contracts for developed, built plots is about 1.500; their value balances around 500 Mio €.

1200 to 1500 contracts are concluded concerning residence and part ownership with a value of about 60 Mio €.

Behind the curtain of this development, over the years there are a lot of different influences changing from year to year. The last ascent of values is reckoned to be based on the fear of people for their financial assets, not coming from a bigger need on the renter's side.

From "concrete gold" people expect to gain the best long-term maintenance of value.

Market economy is a contract economy. The exchange of property only means the top of the economic contracts in the real estate economy. The broad base of contracts is the renting contracts.

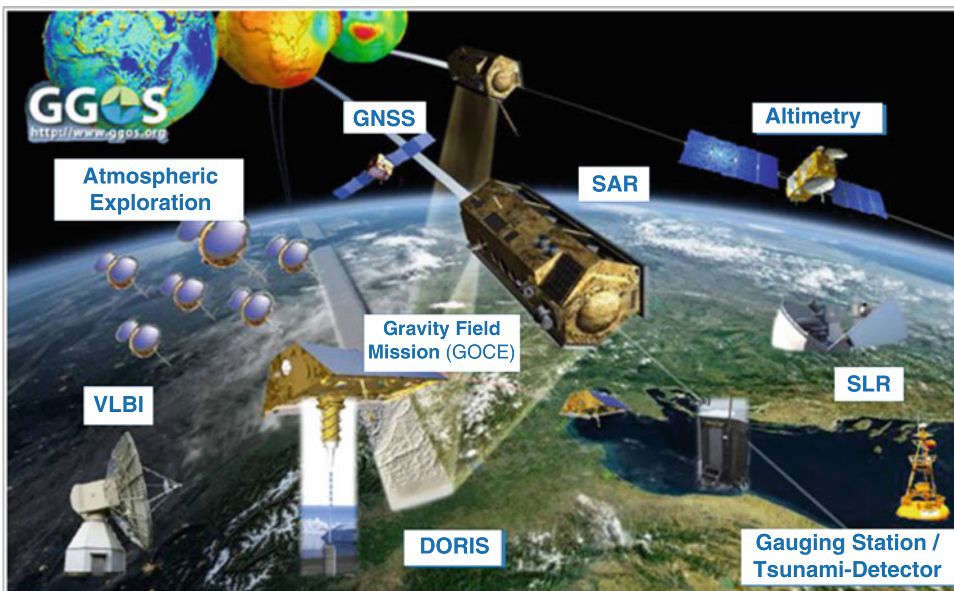
An exchange by contract of about one billion € releases work for the service institutions register and cadastral office, notary, broker, and public appointed surveyor of about 50 Mio €. The transaction tax for the community as well costs the parties 5% which means about 50 Mio €.

Geodetic Technology

Geodetic technology is more productive than ever in history (Schuster 1985). The number of sensors and items grew enormously: satellites for GNSS, altimetry, gravity field missions, side-looking airborne radar (SLAR) or synthetic aperture radar (SAR), and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) (Fig. 11). Not to forget the equipment for atmospheric exploration, the Very Long Baseline Interferometry (VLBI) and the tide gauges or even complex systems like tsunami detectors are all being used in the low and medium earth orbits.

With remote sensing a new geodetic technique came up for evaluation of the satellite photography, laser and radar surveys.

New sensor technology presents a variety of technical devices with regional impact like surveying planes or helicopters, steered by IMUs (inertial measurement units) with a variety of aircraft cameras and lasers. Most of the sensors built have a local impact: GNSS Rovers for dm



Geodesy, Fig. 11 Geodetic technology in the earth near space (Courtesy to Prof. Harald Schuh, GFZ Potsdam)

Geodesy, Fig. 12 Geodetic sensors with local perimeter

G

or cm – accuracy, combined positioning and sensor systems for vehicle fleets or personal tracking as well as tachymeters as total stations and laser scanners of different type (Fig. 12).

In the metrological scale, we find many robotics and high-precision measuring instruments (laser tracker ao), necessary for 3D measurements, and in-house laser systems in a local reference system.

The geodetic methods are directed upon maps and scientific results, but also upon data pools consisting of maps and alphanumerical information like property cadastre and surveying and engineering geodesy. They are used for the control of the building process. There is a big variety of procedures, products, and results. As far as they are based on measurements, they are to be distinguished by an implicit or explicit stochastic. For long the maps developed to graphic interactive systems (GIS). The orthophoto as an automatically generated aerial scalable photo is a worldwide proven and successful product.

Unexpectedly the unmanned aerial vehicle (UAV) or shortly drone started its triumphal march into the geodetic technology: equipped with cameras of most different quality or even small lasers, they catch up the objects in the space near to the surface (Haala and Schwieger 2017). The point clouds produced lead to new economic useful results, i.e., for monitoring, quantity or mass calculation, urban models of higher level of detail (LOD), roof landscapes as orthophotos (Fig. 13), a geodetic base for the building information modeling (BIM), or disaster control. Powerful geodetic software modulates the transformation of the point clouds.

Geodetic laser scan or even laser slam methods (without any fixed standpoint) capture new fields of application spirited up by the revolutionary building information models, which need and are able to process Big Data results during the construction process (Fig. 14).

Parallel to mechanical engineering, the quick comparison between the actual status of a construction in progress can be realized with an authentic point cloud by laser slam, which is called “trusted living point cloud” (Fig. 15).

The actual expansion of the sensor technology and the velocity of the data capturing enable the geodesy to do its bit for the “smart construction site.” The cross-linking of sensors, devices, machines, and finally the workforces is beginning (Fig. 16). The result is a seamless communication between sensors and server over the Internet. The sensors deliver the current necessary information for workforce and fleet management and support the accounting of the site’s activities. Complex algorithms and modern mathematical filter methods have to be composed.

Inertial measurement units (IMUs) in connection with great effort in terms of data filtering will play a great role in the future geodetic technology,

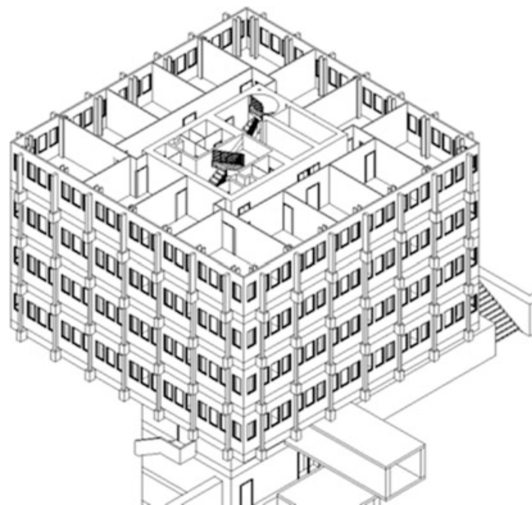
Geodesy, Fig. 13 Roof landscape as an orthophoto (acc. $<\pm 1$ mm) captured by UAV

because they are able to fix the coordinates of a moving object down to [mm] accuracy together with GNSS as a frame. So the coordinates of any moving object can be created coincidently. One scientific challenge of the hour is the high-precision time adjustment of the different moving sensors.

The New Age of Geo-Information

Geodesy means basically the technique, which connects freely selected points with an imaginary line. This technique got social and economic importance with the property cadastre. The border points were surveyed, and the imaginary line between them was the property boundary. It is the borderline between the spheres of activity between neighbors. These borderlines are the most important legal and economic reference frame (Schuster 1997). Most of the authoritative actions and private measures relate to that system. Coordinate systems – as described above – are from that view auxiliary systems, which ease to comprehend the economic life.

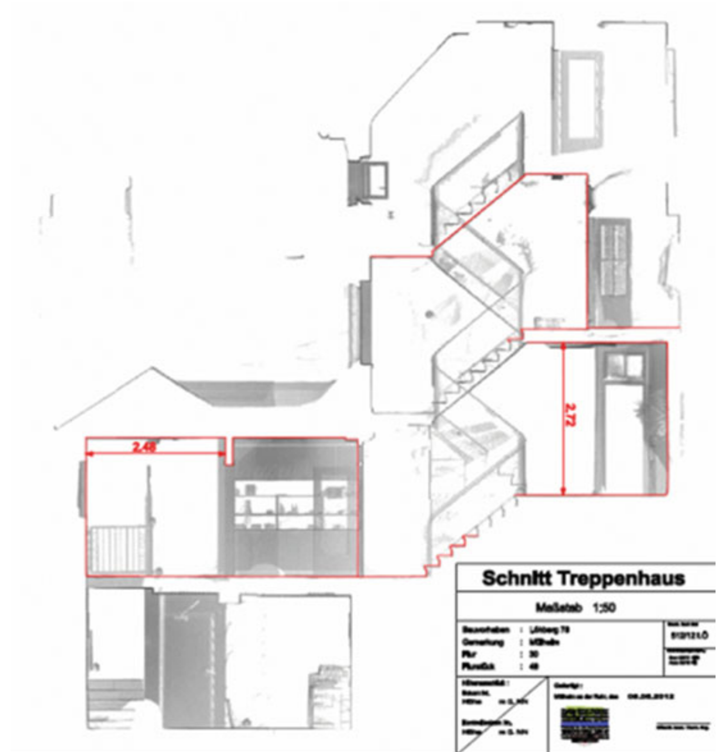
Two hundred years ago, the installation of such a reference system "property borders" needed a huge economic effort of the authority. Today, the sum of new technologies enables the geodetic community to give a coordinate to all points, on which the society wants to imprint its will. Since that procedure happens quickly, the movement of objects or even that of men can be documented in such coordinate systems.



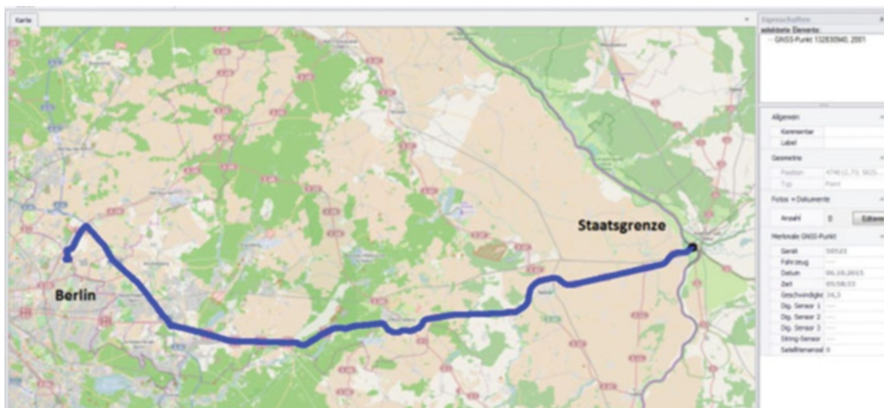
Geodesy, Fig. 14 Building captured by laser slam – base for BIM

Vice versa every coordinate created bears a certain information, the geo-information (3. Geo-Fortschrittsbericht der Bundesregierung 2012).

Seen from an economic viewpoint, the elements of the unnumbered nature and that of human privacy become finite-measurable. This means that water, air, wood, etc. a hundred years



Geodesy, Fig. 15 Trusted living point cloud



Geodesy, Fig. 16 GNSS theft protection system quickly alarms the construction site

ago were unnumbered "free goods"; every human being was allowed to use or pollute them. Since the pollution is measurable, every tree is identifiable and such countable items run short in the mind of the people and economically as well. They even run short in a double sense of the word, physically and economically.

The information – attached to the coordinate – has a value, even when it is outdated. This value has not an eigenvalue in the economic sense; it has only a value in relation to the primary object: the "plot" or the "natural person" or "legal entity."

These newly created huge amount of data need an evaluation to be understandable. Therefore

meta-levels are needed, which means the graphic design in gliding scales and gliding adapted degree of abstraction.

Meta-data as well mean numerical description, the naming of maps and the description of the accuracy of the coordinates or related structures.

The public authorities gain a lot of advantages for their work on future topics like climate, energy, mobility, sustainability, and demographics. The European authorities started early with INSPIRE and Copernicus (former GMES) (Allessandro and Claude 1999). On national and regional level, the European structure got substructures (i.e., GDI-DE) partitioned to responsibilities.

The INSPIRE directive gives the legal frame for that Europe-wide geo-data infrastructure.

Worldwide suppliers of geo-information like Google or Microsoft today seem to be unavoidable, and they disperse in the economy with new business models (Barwinski and Schuster 1988; Schuster 1997).

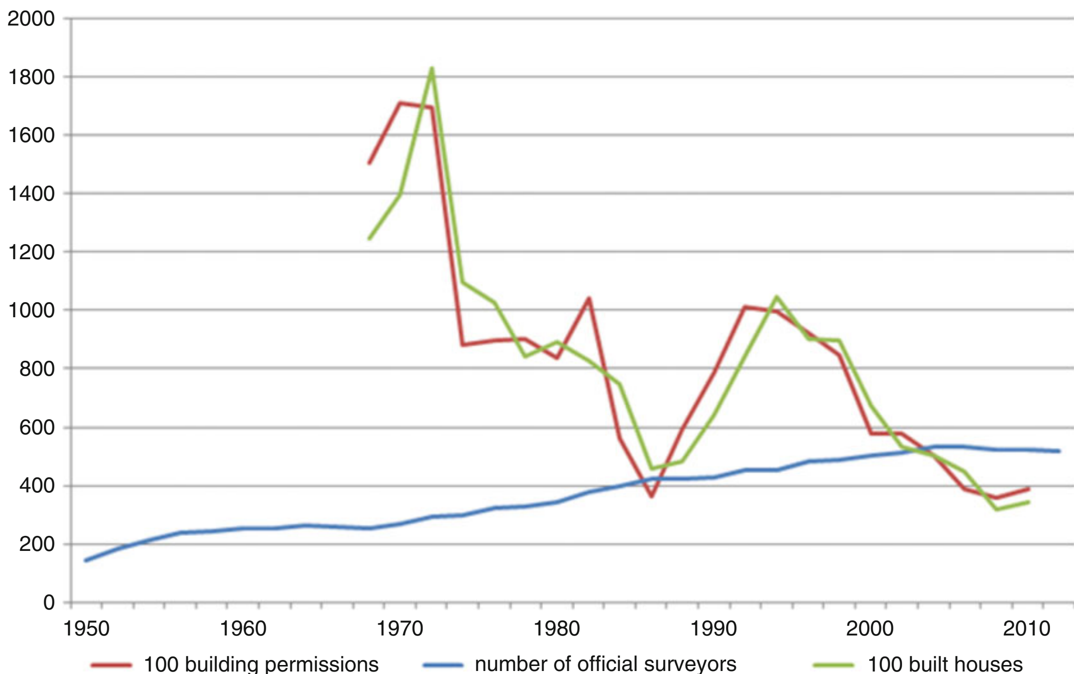
Discussing the naming of this new field of work, international groups decided to change the term to “geomatic” to show that the geodesy has expanded its field.

Geodesy as an Economic Sector

It is comparatively simple to value the costs of an economic sector on the European continent, because the professionals remain for a very high percentage of their professional life in their profession once chosen.

The “Market Report” presented by the Comité de Liaison des Géomètres (CLGE) and Geometer Europas (GE) (Schuster et al. 2003) comes to the conclusion that the volume of costs, deduced from the salary payments to geodetic professionals in the year 2000, is about 24 billion €. This sum comprises those who are busy with the maintenance of the public property system and public maps as public servants, as well as those who are active in the private business life in form of surveying, evaluation, and software (Schuster 2003).

Given that the predominant part of demand in geodesy is due to the building applications, Fig. 17 shows the dramatic changes in demand



Geodesy, Fig. 17 Dramatic changes of supply and demand

since the 1950s and the steadily growing number of supply (number of private offices) in a state like North Rhine-Westphalia with about 18 Mio inhabitants.

The private service structure of the supply answered with the reduction of personnel and an furthermore improved its offering with better quality of comprehensive services on the basis of the extending public regulations; the public structure passed several waves of reorganization but benefitted from the extending relevant legislation.

So, the demand for geodetic supply is relevant to the legal situation of land as a whole. The differences between countries can be huge.

In a developed country like the Netherlands, the geodetic community speaks about "crowdsourcing," which means that the owners themselves are able to identify the property boundaries between their parcels on the basis of a well-explained cadastral information of the public authority.

The Internet now has brought up a new situation, where some hundred experts, i.e., from Google and Bing, can theoretically reach the whole mankind with their products. The funds for this investment and its maintenance come mostly from sources outside of the professional sphere (i.e., marketing). The underlying business model couples different sectors, which is not unusual in the economy (Schuster 2004).

Whereas the old business model of the geodesy was relatively closed and strongly regulated and could be broken down into

- Public mapping
- Maintenance of cadastre
- Cadastral surveys
- Sequent control of building geometry,

the business field expanded its supply strongly by the vast new field of sensors, the practically unlimited possibility for the geo-referencing of any object, and the use of the Internet. The enlargement of the supply is financed in the middle-class economy by the old business, big public investments, and the hope, which was induced by the technological progress. The demand is lagging behind, which means generally low prices for the services.

The situation in various European countries is different, and it is scarcely straightforward because of the degree of regulation, not to mention the clearness in detail. Beside such difficulties one can say that the old geodetic business model was pillared by:

- Ground tax
- Transaction traffic
- Settlements
- Military needs
- Sequent control of building geometry

Nowadays is upcoming a growing income for the use of most different geo-information. One quickly spreading example is the area cadastre. With this GIS the communities are able to split the fees for percolation water and sewage. New services are created. With the cadastre of trees, of greens, and of public eases, they help to manage the real estates and have an organizational added value.

Many people look forward to a growing activity in geodesy. But the answer to the question is not yet clear, until new "biting" business models and broader economic trends will show up. Undoubtedly, they will deform the old economic basis.

References

Allessandro A, Claude L (1999) (Europäische Kommission) (eds) Spatial reference systems for Europe. A joint initiative of MEGRIN and the Space Application Institute

Barwinski K, Schuster O (1988) Landinformationssystem und Freie Berufe. BDVI Schriftenreihe des Bundes der Öffentlich bestellten Vermessungsingenieure, vol 1

Bundesministerium des Innern: 3. Geo-Fortschrittsbericht der Bundesregierung Stand Okt (2012)

Haala N, Schwieger V (2017) UAV – Anforderungen und Möglichkeiten. DVW Schriftenreihe des Deutschen Vereins für Vermessungswesens, vol 86

Kahmen H (2005) Angewandte Geodäsie: Vermessungskunde, 20th edn. Gruyter, Walter de GmbH, Berlin. ISBN 3-11-018464-8

Kerl V, Schuster O (2003) Necessary outline conditions for property financing through covered bonds and

mortgage banks. GEOhaus GbR, Berlin/Mülheim an der Ruhr
Ledersteger K (1969) Astronomische und physikalische Geodäsie, Handbuch der Vermessungskunde, Bd 5, 10. Aufl. Metzler, Stuttgart
Schuster O (1981) Der Berufstand des Vermessungsingenieurs und seine Vorläufer, Schriftenreihe des Vermessungstechnischen Museums. Konrad Wittwer Verlag, Dortmund, S 83–101
Schuster O (1985) Satellitennutzende Vermessung – Beginn einer neuen Ära. BDVI-FORUM Zeitschrift des Bundes der Öffentlich bestellten Vermessungsingenieure. p 185
Schuster O (1997) Freie Berufe und Kultur – Schlaglichter. BDVI – Schriftenreihe, vol 6
Schuster O (2003) Geodäsie in der Wirtschaft „FuB“ -Flächenmanagement und Bodenordnung Heft 2
Schuster O (2004) Geodesy in the German economy. Eur J Law Econ 17(2):191–208
Schuster O (2005) Spatial data and globalisation, Atlantic Think Tank VI. Massachusetts Institute of Technology, Cambridge, MA
Schuster O, Ouranos E, Busch M, Höflinger E(†) (2003) Market report – report about the Market of Surveying in Europe, Editors: Comité de Liaison des Géomètres (CLGE) und GEOMETER EUROPAS (GE)
Seitz M, Bloßfeld M, Angermann D, Schmid R, Gerstl M, Seitz F (2016) The new DGFI-TUM realization of the ITRS: DTRF2014 (data). Deutsches Geodätisches Forschungsinstitut, Munich. https://doi.org/10.1594/PANGAEA.864046
Sella GF, Dixon TH, Mao A (2011) REVEL: a model for recent plate velocities from space geodesy. J Phys Res 107(B4):ETG 11-1–ETG11-30
Torge W (2009) Geschichte der Geodäsie in Deutschland, 2. Aufl. De Gruyter, Berlin. ISBN 978-3-11-020719-4

Geographical Indications

Marco Ricolfi
Department of Legal Studies, Department of Law, Turin University, Turin, Italy

Definition

The issues at hand. Geographical indications (GIs) are signs (or symbols) used in connection with the sale of goods which convey an association, direct or indirect, with a place or location, which may be relevant for certain features of the good in question.

Rules concerning GIs may be found both at the municipal (or domestic) level and in international agreements. The international framework for dealing with GIs is to be found in the 1883 Paris Convention, in the 1891 Madrid Arrangement and in the 1958 Lisbon agreement. After the adoption by the European Union of the *sui generis* regime for GIs and designations of origin in 1992, a multilateral equilibrium was found in the 1994 TRIPs agreement.

When the same geographical symbol is used simultaneously by many businesses, a number of issues arise which may be dealt with in several ways. The different alternatives are discussed hereafter along with the reasons which may dictate the choice among them. They all rely, however, on a common assumption: that GIs should, at least in principle, not be appropriated by one single business at the expense of all its rivals. Therefore, we first deal with this assumption, which is shared by most legal systems.

Geographical Symbols and Individual Trademarks

Most legal systems prevent a single business to register as a trademark a geographical symbol (or to otherwise obtain exclusive legal rights over it). The reason for this prohibition is straightforward and has two prongs. Let us assume that the business trying to appropriate the geographical symbol is in fact established in the geographical area to which the symbol alludes to. If this is the case, and the geographical origin is responsible for features which are relevant for consumers' choices, then the grant of a trademark registration would generate an unwarranted competitive advantage for the benefit of one business to the exclusion and detriment of all its competitors. The matter gets even worse if we consider the second prong: if the business holding the trademark is not even located in the geographical area to which the symbol alludes and, therefore, the goods originating from it do not possess the corresponding features, then the public is deceived.

Usually, these difficulties are tackled by four related rules:

- There is a bar on the registration of signs which convey an association between features of the good and a given place (in US law: primarily geographically descriptive symbols).
- Should in fact the goods bearing the trademark in violation of the previous rule also come from an area different from the one designated by the symbol, this would also amount to prohibited deceptive behavior.
- However, the two rules above do not prevent the registration of a geographical symbol which, on its face, has no association whatsoever with features of the goods on which it is affixed; Montblanc fountain pens is a good example of this situation.
- Also, in some – but not in all – legal systems registrability may be obtained through the acquisition of a "secondary meaning," which, in this case, indicates that the relevant public no longer associates the indication with a geographical place but with a business origin which has obtained recognition in the marketplace.

One Geographical Symbol, Many Users: The Legal Options

All the difficulties associated with the use of a geographical symbol, discussed in the previous paragraph, fade away when the geographical symbol is used simultaneously by several businesses, all located in the area which is relevant for the characteristics of the goods. This is an occurrence which is more frequent in connection with agricultural goods, such as wines, cheese, fruit, and the like, but may occasionally be relevant also for manufactured goods which embody handicraft traditions, such as pottery and fabrics (Cuccia and Santagata 2004). This does not mean that the simultaneous use of the same geographical indication by different businesses is unproblematic. To the contrary, this situation may generate a number of difficulties, which, while different from the ones arising from the use by a single business, still require the adoption of appropriate legal devices to avoid the emergence of recurring shortcomings, which include consumer deception and unwarranted competitive advantage, again, but are not limited to them.

Three are the main options available to deal with the issues arising by the use of geographical symbols by multiple businesses, which may consist in the setting up of property rights, in the resort to liability rules, and in the adoption of a *sui generis* regime. While the choice between these different options ultimately rests on international considerations, rather than on the comparative assessment of the economic costs and benefits of the different approaches, for analytical reasons we will first describe the different options available at the domestic level in an economic analysis of law (EAL) perspective, reserving for a later stage consideration of the international factors which determine the choice.

G

The Domestic Level: Property Rights

It is possible to set up property rights enabling multiple businesses to use the same geographical symbol in a coordinated and nondeceptive manner in two different ways: collective trademarks and certification marks. Both are known in common law systems as well as in civil law countries. However, the latter tend to favor collective trademarks, the former certification marks (Belson 2002; on the status of collective and certification trademarks in the TRIPs Agreement see Pires de Carvalho 2006; on the corresponding notions Spada 1996). In either case, the geographical sign is registered as a trademark; as a result, any use by an unauthorized third party of a sign which is identical or similar for the goods indicated in the certificate is automatically considered as an infringement.

What distinguishes collective and certification trademarks is that in the former, the holder of the trademark is an entity (a corporation, a consortium, an association) and the users are members of the same; in the latter, the holder of the trademark is an independent third party, with no links with the users, which is legally prevented from engaging in the business for which the trademark is registered. This third party is contractually

bound to certify that users comply with the applicable standards, as agreed in an appropriate document (the "specification").

In short, collective trademarks are membership-based; certification trademarks are contract-based. All the differences in the rules concerning collective and certification trademarks flow directly from this distinguishing feature.

Collective Trademarks

Each user obtains authority from the entity which holds the trademark to use it in conformity with the "specification," which is adopted and submitted with the application for registration. In principle, only businesses which are members of the entity holding the registration are allowed to use the collective trademark. On the one hand, this rule is consistent with the more communitarian underpinnings of collective trademarks: only businesses which operate in the geographical area which is relevant for the features of the goods or otherwise share a characteristic which is signaled by the collective trademark are allowed to resort to it. On the other hand, making membership of the trademark-holding entity a precondition for use entails the possibility of abuse. More specifically, the risk is recurring that businesses complying with all the requirements are denied the opportunity to become members of the collective trademark holding entity and therefore to obtain a license to use it. Legal systems tackle this risk in several ways (Sarti 2011; Spada 1996). In some systems, an "open door" policy is mandated: applications for registration of collective trademarks are granted only on condition that the by-laws of the applicant entity provide that each business meeting the requirements is allowed to become a member and to obtain a license to use. In other systems, even non-members are entitled by law to use the geographical indication corresponding to the collective trademark, provided the sign is used as a descriptor of geographical origin rather than as trademark. Legal systems which do not adopt these approaches may resort to antitrust laws to make sure that the denial of admission to membership of a business meeting all requirements is prevented.

Certification Trademarks

The entity holding the certification trademark may be public or private and may come in any legal form. What is of essence is that its relationship with users is contract-based: compliance with rules set in the specification entitles to use of the trademark; no discrimination is allowed in the certification. This is the rational basis of the rule whereby the certifying entity may not be in the business sector for which it gives the certification.

The Domestic Level: Tort

Businesses located in a certain area which is relevant for some of the features of the goods they supply have not only the option to obtain a property right in the geographical indication, but, as an alternative, they may also resort to tort law. This is particularly so in legal systems which give protection against unfair competition. However, the scope of protection afforded under liability rules is more limited than the one based on property rules. A predicate required before a business operating in the relevant area may claim protection against use by a third party to associate with the location is that the geographical sign has already reached a certain degree of recognition or reputation. It should be noted in this connection that, when protection is sought at the international level, recognition in the country of destination (import), not of origin (export), is controlling. The businesses operating in the relevant area must also prove that the third party resorting to the corresponding geographical symbol is taking advantage of a reputation it has not earned ("free ride"), to the detriment of lawful users ("injury") and, additionally, that the public is likely to be deceived. This requirement entails another limitation in protection; disclaimers, such as corrective additions on the label (e.g., Roquefort, produced in Ontario, Canada) or qualifying language (e.g., Roquefort-type), may avert liability.

A Comparative Assessment

There is not much literature comparing the costs and benefits of protection of GIs under property

rules as opposed to liability rules. This is rather unsurprising, as the protection of GIs is not prevailing in the Anglo-Saxon world (Faulhaber 2005), and European scholars are wary of treading in waters not tested by American literature (Ricolfi 2009a). However, in a welfare maximization perspective, we may quite safely assume that the property and liability rules are effective in a roughly equivalent way in pursuing the goal of lowering consumer search costs. As to the other function of trademarks, to provide an incentive to invest in quality (or, rather, a combination of quality and price which meets consumers' desires), it is submitted that, on the one hand, property rules tend to be superior. Resort to a standard set by the relevant community in a formal and transparent way, the specification, is likely to avert resort to free-riding better than a reference to more nebulous "best practices" to be established during litigation. However, the costs of setting up a collective trademark registration system are likely to be substantial (Van Caenegem 2004). On the other hand, assuming that adjustment over time is a desirable feature for GIs, where standards may change over time, it is arguable that tort-based systems are superior to property-based systems, as they tend to be more flexible. This is so because they rely on the understanding associated with a specific geographical symbol at the time a controversy arises rather than on the notion codified once and for all in a specification.

The Sui Generis Regime of Protection of Geographical Indications

In the quest for protection of geographical symbols, there is a third alternative which tops the property and liability rules-based systems just sketched out: the *sui generis* regime adopted back in 1992 by the EU, which is remarkable in itself and as an (alleged) blueprint for the multilateral provisions adopted by Artt. 22 ff. TRIPs. The current version of the EU regime is to be found in reg. no. 1151/2012. All in all, this regime is more similar to property than to liability rules-based systems, even though it deviates from them in that there is no person or entity which may be described as the owner or holder of a sign. The collective monopoly, bestowed on all businesses which can show their location in the relevant area and compliance with the requirements set by the specification, is the result of a public law proceeding initiated at the local level and completed by the grant of title by the EU Commission. What is provided by this set of rules is a *sui generis* regime, characterized by low access requirements and strong levels of protection. This feature is not very usual in intellectual property law, where, by and large, the level of protection is commensurate to the access requirement. The deviation shown by the EU *sui generis* regime is accounted for by a clear political choice intended to give a high level of protection to agricultural communities and to their local traditions in growing and producing foodstuffs.

Three components of the regime stand out.

1. As far as *access requirements* are concerned, the regime, Art. 5 of reg. no. 1151/12, applies to so-called protected designation of origin (PDO) which, as in the prior international framework, refers to signs identifying products whose quality or characteristics are essentially or exclusively due to a particular geographical environment with its inherent natural and human factors (the so-called *milieu*). The protection extends, however, also to "protected geographical indications" (PGI), the requirement for which are remarkably lower as they depend not on objective criteria (the *milieu*) but on the subjective factor of the reputation and even admit to some relaxation as to the localization of the production steps;
2. While in theory "generic terms shall not be registered" as PDOs or PGIs, Art. 6 of reg. no. 1151/12, in practice even descriptive terms, like "Feta," for the characteristic cheese of Greek origin, have been found to be registrable even though they had been in common usage outside Greece for long time (the ruling by the European Court of Justice was adopted in 2005). Under Art. 13 PDOs and PGIs, once registered, are no longer subject to genericization.
3. Registration as PDO or PGI confers an exclusivity over the sign entailing very strong

G

protection. While the protection for ordinary trademarks is granted only against the risk of confusion, here the grant extends to usage which falls short of such a risk (e.g., mere "evocation") and does not take into account the presence of disclaimers which may prevent consumers' confusion (Art. 13 of reg. no. 1151/12).

Originally the *sui generis* regime was thought of as a tool to facilitate the replacement of member States rules by a single EU protection title. However, Art. 11 of the current regulation opens up access to this EU title also to non-EU entities, in furtherance of the TRIPs-mandated prohibition of discrimination against non-EU nationals.

The comparative assessment carried out above may be extended to the *sui generis* regime. Overall, the regime is closer to property than to tort: access requires registration; once registration is obtained, no proof of deception of the public or of injury to lawful users is required. What is remarkable in the regime is the political nature of the process which leads to the registration. For example: one single Greek island may register a dozen GIs for olives and oil, not so much because there is a business rationale for this proliferation but because local politicians and communities choose to engage in an exercise where most of the costs are borne by the general taxpayer. This is hardly surprising: a preferential treatment for agricultural communities is at the basis of the French-German alliance which has been underpinning the design of the EU from its beginnings.

The International Level

One might assume that lawmakers are in a position to choose among the options described in above in order to maximize the welfare of their constituency and that they accordingly proceed to select the regime which minimizes costs and maximizes the benefits. In real life, lawmakers have another, overriding variable to consider: the position in international trade of the political entity they are legislating for. In this connection, two considerations should be factored in. *First*, the choice has a very strong mercantilist component: usually countries tend to support exports and discourage imports; and this applies also to goods bearing geographical indications. In this regard, producers' interests may prevail over consumers', as predicted by collective action theory. *Second*, goods originating from a given country (country A) may derive competitive advantage (or disadvantage) not so much from rules adopted in country A as from rules applicable in the export markets countries B, C, . . . N. As a result, it is submitted that, while EAL is a powerful analytical tool to examine the competitive consequences of the different options, it is public choice theory which ultimately accounts for the choice among the various options.

This is the backdrop against which the international framework of norms concerning GIs has to be assessed. In doing so, one should never forget that here we are dealing with power structures rather than with economic optimality.

The 1891 Madrid Agreement: The Tort-Based Approach The 1883 Paris Convention, which is considered – along with the 1886 Berne Convention for copyright – as one of the two pillars of international intellectual property protection, only mentioned geographical indications without adopting specific rules concerning them. As a follow-up to it, the 1891 Madrid Agreement was adopted, which visualized protection of GIs as an issue of unfair competition. Access requirements for protection are low. Also simple (and indirect) indications of geographical origin are protected; therefore no specific link between the features and the place of origin is required. Correspondingly, protection is weak: it is confined to cases of actual fraud; even more to the point, the relevant recognition and reputation of the goods designated by a geographical indication has to be established on the sales market, not in the country of origin. In this perspective, the Madrid rules are acceptable to countries which tend to import, rather than to export, goods bearing GIs.

The 1958 Lisbon Agreement for the Protection of Appellations of Origin The countries which

tend to export, rather than to import, goods bearing a geographical indication pushed in the opposite direction half a century later, finally adopting the Lisbon Agreement in 1958. This convention presents features which are the obverse of the ones observed in the Madrid Agreement. The access requirements for protection are high: only signs identifying products, the quality or characteristics of which are exclusively due to the *milieu,* are eligible and their protection is contingent on registration in appropriate national registers. The scope of protection is correspondingly high: it extends to usurpation and imitation; infringement is not ruled out by disclaimers or rectifying language (e.g., "like" or "type"); and, most importantly, it has an extraterritorial dimension, in that, once the appellation of origin is registered in one State party to the Agreement, protection extends to all the other States automatically, without the need to establish either recognition or reputation on the sales market.

The problem with the Lisbon Agreement is that it only includes like-minded countries. The 27 States which are parties to it are notable for the importance and quality of their agricultural production, which makes them export-, rather than import-oriented. For their own part, import-oriented jurisdictions are loath to join an agreement which would increase the level of protection of foreign productions to the detriment of local businesses. In 2015, the Geneva Act has tried to take care of these difficulties, lowering the access requirements to the system. Whether this move will in fact enable the participant countries to reach the required critical mass remains to be seen.

A Half-Way House: The GI Provisions in TRIPs

EU officials like the idea that the TRIPs provisions concerning GIs (Artt. 22–24) follow the blueprint of the European *sui generis regime* (Gervais 1998). This is to a very large extent a delusion. While a regime resembling the EU approach is envisaged for wines and spirits (Art. 23), all other GIs are dealt with as a tort or unfair competition issue: no protection is available without deception of the public (Art. 22). As a result, the EU has placed itself into a situation where it is under the obligation to extend the benefits of its generous *sui generis* regime to all businesses which are resident in a country which is a TRIPs member (Artt. 3–4) while its own businesses cannot claim a protection going beyond tort law for European GIs abroad, which was already available to them under the 1891 Madrid Agreement all the time. The EU appears all the more misguided as it tries to lure developing countries to follow it into a bargain in which it is a clear loser (Ricolfi 2009b).

Future Directions: Conflicting Agendas in GI Protection

There is no doubt that the EU and the Anglo-Saxon world are pushing in opposite directions as far as GIs are concerned (O'Connor 2004). The latter favors GI rules which avoid competitive disadvantage deriving from location for their powerful agro-businesses. The former extends into the twenty-first century the traditional pro-agricultural communities attitude inherited from the French-German arrangements at the basis of the European architecture. What we do not know as yet is how this divergence will fare when brought into the much wider compass of global competition. Initially, the rival views have played out in the bilateral and regional free-trade agreements (Rangel-Ortiz 2014). Until the US elections in 2016, it would have seemed that the American approach was taking the upper hand also in international trade negotiations, such as the Trans Pacific Partnership (TPP) and the Trans-Atlantic Trade and Investment Partnership (TTIP). However, the situation is very much in a state of flux now, as a result of the repudiation of these treatises by the Trump presidency.

We may wonder what role was – and is being – played by Economic Analysis of Law in the process. Perhaps unsurprisingly, EAL-based tools may contribute to understanding what are the costs and benefits of the different alternatives at hand, but as the choices are being taken at the super-national, rather than domestic level, they seem to be suggested much more by power dynamics rather than by the pursuit of welfare optimization.

References

Belson J (2002) Certification marks, special report. Sweet & Maxwell, London
Cuccia T, Santagata W (2004) Collective property rights for economic development: the case of the ceramic cultural district in Caltagirone, Sicily. In: Colombatto E (ed) The Elgar companion to the economics of property rights. Edward Elgar, Cheltenham, 473 ff
Faulhaber LV (2005) Cured meat and Idaho potatoes: a comparative analysis of European and American protection and enforcement of geographic indications of foodstuffs. Columbia J Eur Law 11:623 ff
Gervais D (1998) The TRIPs Agreement. Drafting history and analysis. Sweet & Maxwell, London, p 119
O'Connor B (2004) The legal protection of geographical indications. Intellect Prop Q 1:35
Pires de Carvalho N (2006) The TRIPs regime of trademarks and designs. Wolters-Kluwer, The Hague, 214 ff. and 369 ff
Rangel-Ortiz H (2014) Patent and trademark rights in commercial agreements entered by the United States with Latin American nations in the first decade of the twenty-first century: Divide et vinces. In: Ghidini G, Peritz JR, Ricolfi M (eds) TRIPs and developing countries. Towards a new IP world order. Edward Elgar, Cheltenham, 72 ff
Ricolfi M (2009a) Geographical symbols in intellectual property law: the policy options. In: Festschrift für Ulrich Loewenheim zum 75. Geburstag, Schutz von Kreativität und Wettbewerb. Verlag Beck, Munich, 231 ff
Ricolfi M (2009b) Is the European GI policy in need of rethinking? IIC 40:123 f
Sarti D (2011) Il marchio collettivo. In: Ubertazzi LC (ed) La proprietà intellettuale. In: Trattato di diritto privato dell'Unione europea directed by G. Ajani and G.A. Benacchio, vol XI. Giappichelli, Torino, 131 ff
Spada P (1996) Il marchio collettivo "privato" tra distinzione e certificazione. In: Scritti in onore di G. Minervini, vol II, Impresa, concorrenza, procedure concorsuali. Morano, Napoli, 475 ff
Van Caenegem W (2004) Registered GIs: intellectual property, agricultural policy and international trade. Eur Intellect Prop Rev 26:170 ff

Further Reading

Auteri P (2010) Indicazioni geografiche, disciplina delle pratiche commerciali scorrette e concorrenza sleale. In: Studi in memoria di Paola A.E. Frassi. Giuffrè, Milano, 23 ss
Blakeney M (2014) The protection of geographical indications. Edward Elgar, Cheltenham
Echols MA (2003) Geographical indications for food, TRIPs and the Doha development agenda. J Afr Law 47:199 ff
Gangjee D (2012) Relocating the law of geographical indications. Cambridge University Press, Cambridge
Goebel B, Groeschl M (2014) The long road to resolving conflicts between trademarks and geographical indications. TMR 104:829 ff
Knaak R (2015) Geographical indications and their relations with trade marks in EU law. ICC 46:843 ff
Libertini M (2010) L'informazione sull'origine dei prodotti nella disciplina comunitaria. Riv Dirit Ind I:289 ff
Mantrov V (2014) EU law on indications of geographical origin: theory and practice. Springer, Cham
Ricolfi M (2015) Trattato dei marchi. Diritto europeo e nazionale, vol 2. Giappichelli, Torino, 1754 ff
Sarti D (2009) Segni e garanzie di qualità. In: Ubertazzi B, Muñiz Espada E (eds) Le indicazioni di qualità degli alimenti. Giuffrè, Milano, 113 ss
Sironi GE (2014) Marchi collettivi. In: Scuffi M, Franzosi M (eds) Diritto industriale italiano, vol 1, Diritto sostanziale. CEDAM, Padova, 312 ff
Tosato A (2013) The protection of traditional foods in the EU: traditional specialties guaranteed. Eur Law J 19:545 ff

German Law System

Sonja Elisabeth Haberl
University of Ferrara, Ferrara, Italy

Abstract

The following entry provides an overview of some elected aspects of the German law system. From different points of view, the German system has been deeply influenced by the ordoliberal ideas developed within the Freiburg School in the early 1930s of the twentieth century. One of the core ordoliberal concepts that have to be discussed within the constitutional framework is that of a social market economy. Social market economy became the interpretive framework for the economic and social order of Western Germany in the aftermath of World War II, and today, it represents not only a key concept at national level but also within the European Union. No less important is the role of the so-called Private Law Society, another key concept of ordoliberal thinking. Its main elements are clearly reflected in the Bürgerliches Gesetzbuch (BGB) of 1900 which is based on the idea of the citizen as a *homo oeconomicus*. Notwithstanding its traditional approach – libertarian, unsocial, and

individualistic – the BGB, a child of the abstract conceptualism of the Pandectist school, has been able to survive till today. This is the merit of judge-made law and, in particular, of the theory of the indirect horizontal effect of fundamental rights in relations governed by private law. In recent times, the BGB has even assumed a highly visible role as a possible model within the harmonization of European contract law. The entry finishes with a description of the German system of legal education, a state-oriented and judge-centered bureaucratic model which is still embedded in the model of a "uniform jurist," the so-called *Einheitsjurist*.

The Constitutional Framework

The memory of the collapse of the ill-functioning, weak, and helpless Weimar democracy which facilitated the slide into a totalitarian dictatorship has profoundly influenced the framers elaborating the post-1945 constitutional order of a West German state (Kielmansegg 1990). The document which should offer appropriate guarantees for the development of a solid democracy, able to defend itself against its enemies, was elaborated by a Parliamentary Council, composed of 65 members elected by the state parliaments, and approved by the *Länder* rather than by popular referendum. It entered into force on 23 May 1949 and was named *Grundgesetz* (Basic Law) and not *Verfassung* (Constitution) in order to underline its provisional character while waiting for unification. Despite this, the Basic Law outgrew its temporary character by means of the accession of the five East German *Länder* to the FRG in October 1990, and with only a few provisions revised and others inserted, it became an all-German constitution within the new Berlin Republic.

The Discussions About the Basic Law's Economic Order: Social State, Social Market Economy, and Ordoliberalism

Unlike the Weimar Constitution, the Basic Law does not contain any explicit reference to social rights. The only indirect references are included in Art. 20, which defines Germany a "democratic and *social* federal state," and in Art. 28 which requires the constitutional order in the *Länder* to conform "to the principles of a democratic and *social* state governed by the rule of law" (the so-called principle of homogeneity). The reasons for the mere affirmation of the social state principle are closely connected to the highly conflicting positions between the constitution-shaping political forces with regard to the economic order which the Grundgesetz should refer to. While the Social Democratic Party tried to promote the concept of economic democracy, based on a form of planned economy, workers' participation in management, and the nationalization of important economic interests (Bommarius 2011), the Christian Democratic Party argued in favor of a system of neoliberal democracy, in which state intervention should be aimed at creating the necessary conditions for the functionality of the market mechanism. Since the various political forces involved in the constitution-making process were not able to reach a consensus on this issue, they set aside the prescription of a specific economic model in the Grundgesetz. As the Federal Constitutional Court later stated, the Basic Law is "neutral" as regards the economic order (4 BVerfGE 7 (1954)). Instead, they contented themselves with reference to the principle of the social state, thus giving rise to the paradoxical fact that the Basic Law ascribes a fundamental rank to "the social," yet without defining it more precisely by means of specific social rights.

After a short time, a new key word was coined through which the task of the German State to perform as a "social state" should be fulfilled (Zacher 1987): the ideas underlying the concept of a social market economy are theoretically and ideologically deeply rooted in the ordoliberal ideas developed within the Freiburg School during the Nazi dictatorship. In harsh opposition to the Weimar party and intervention state (Stolleis 2002), its proponents – among others the economists Eucken, Rüstow, Röpke, and the lawyer Böhm – were united in the idea of militating against social and political pluralism and arguing in favor of a "strong state," called to efficiently

manage the economy by building and enforcing a legal regime representing an *ordo* intrinsic to economic life (Joerges and Roedl 2004). The ordoliberals thus propagated the so-called third way as the proper alternative between laissez-faire liberalism and collectivist forms of political economy, calling for the establishment of a system of undistorted competition in order to enhance the functioning of the economic order.

It was Alfred Müller-Armack who substantiated the ordoliberal ideas in the enticing slogan social market economy and transformed it into policy together with Ludwig Erhard under the chancellorship of Konrad Adenauer. Social market economy refers to a strategy based on voluntary market transactions with undistorted competition and private law mechanisms as its fundamental elements, to be complemented only by a certain kind of state intervention: contrary to what one could think, in fact, the meaning of the attribute "social" does not refer to a policy of social justice associated with a welfare state. Müller-Armack, by arguing for the "total mobilisation of all forces" (Müller Armack 1933), to be achieved by enabling individuals as self-responsible and self-determined entrepreneurs (Bonefeld 2012; Somma 2013), primarily referred to the efficiency of a market economy, i.e., to the fact that the latter generates economic growth and wealth, thus directly and automatically bringing about social achievements (Joerges and Roedl 2004). In addition, he required "a system of social and societal measures" as social peacekeeping tools which however had to be *marktkonform*, i.e., consistent with the competitive order (Müller-Armack 1966).

Social market economy has become a core concept of the foundational period of the FRG, whose success story in the postwar period is frequently associated with it. With its insertion, in 1990, into the treaty establishing a Monetary, Economic, and Social Union between the FRG and the GDR, the term achieved the rank of a legal norm for the first time. It is also included in the Treaty of the European Union (Art. 3) which sets among its objectives that of a "highly competitive social market economy."

The System of Fundamental Rights and the Bundesverfassungsgericht as the Guardian of the Grundgesetz

The outstanding importance of the fundamental rights included in the *Grundgesetz* is emphasized by their location – they are all placed at the head of the document (Arts. 1–19), with the guarantee of human dignity at its core. In the words of the Federal Constitutional Court (Bundesverfassungsgericht), the paramount importance of Article 1 par. 1 – which declares human dignity as inviolable and provides for its obligatory respect and protection by all state authorities – can be explained only by the historical experience and spiritual-moral confrontation with the previous system of National Socialism. In order to avoid another abyss, the Grundgesetz has constructed a "value-bound order in which the individual person and his/her dignity is placed at the center of all its rules and regulations" (39 BVerfGE 1 (1975)). Art. 1 par. 1 thus expresses the highest value of the *Grundgesetz*, which not only has an effect on the following fundamental rights but also informs the substance and spirit of the entire document (Häberle 1987).

The whole system and structure for the protection of fundamental rights aim to correct the mistakes of the past. While the fundamental rights included in the Weimar Constitution were only declaratory and thus not judicially enforceable, the Basic Law's fundamental rights are transformed in subjective rights by Art. 1 par. 3 which states that they "shall bind the legislature, the executive and the judiciary as directly applicable law." In addition and also directed at remedying the deficiencies of the Weimar Constitution, the *Grundgesetz* contains precise rules and limits which have to be respected in the context of the restriction of a fundamental right, whose "essential core" (*Wesensgehalt*) may in no case be infringed (Art. 19).

On the other hand, the Grundgesetz provides for the forfeiture of certain fundamental rights, should they be abused "in order to combat the free democratic basic order" (Art. 18). Besides the possibility of a party ban (Art. 21) and the

prohibition of associations (Art. 9), the forfeiture of basic rights is considered to be one of the core elements regarding the conception of democracy in the Basic Law, qualified as a "militant democracy" (*streitbare oder wehrhafte Demokratie*), which "expects its citizens to defend the free democratic basic order and does not accept the misuse of fundamental rights aiming to undermine it" (28 BVerfGE 36 (1970)). In order to further strengthen this concept, the Parliamentary Council opted for the textual anchoring of a so-called eternity clause (*Ewigkeitsklausel*) according to which certain core elements of the Constitution are unamendable (Art. 79). The federal state principle, human dignity, and the basic principles of state order mentioned in Art. 20 belong to the elements which represent the identity of the Basic Law and are thus immune to any constitutional revision.

The supreme guardian of the Constitution is the Bundesverfassungsgericht (BVerfG). It was established in 1951 and is vested with extraordinary powers. It has not only the competence for constitutional disputes between federal organs and between the federation and the *Länder* but also the right to control the constitutionality of laws and to deal with individual constitutional complaints which can be initiated by a person who alleges to be negatively affected in his/her fundamental rights by an action of the public authority. The proceedings for controlling the constitutionality of laws can either be initiated by an ordinary court (concrete judicial review) or upon application of the Federal Government, a *Land* government, or one fourth of the members of the *Bundestag* (abstract judicial review). In addition to these competences, the BVerfG also rules "in the other instances provided for in this Basic Law" (Art. 93). It is the competent organ for declaring the forfeiture of basic rights (Art. 18), for cases of impeachment of federal judges (Art. 98) and of the federal president (Art. 61). Its function as a guardian of the Constitution perhaps finds its most evident expression in Art. 21 according to which the BVerfG shall rule on the question of the unconstitutionality of political parties.

The Federal System and the Division of Powers

The federation established after reunification in 1990 consists of 16 *Länder*, and each of them has its own constitution upon which the state institutions are based. Although both the *Bund* and the *Länder* are vested with the three branches of public power, these are not separate and distinct from each other. On the contrary, Germany's system of cooperative federalism gives a significant example of what Fritz Scharpf termed *Politikverflechtung* (Scharpf 2009), referring to the complex interrelationship, interdependencies, and overlappings between the competences of the federation and the states. The Federalism Reform I of 2006 aims to disentangle these interdependencies, on the one hand by more precisely delimiting the allocation of legislative competences between the federal and the state level and on the other by reducing the risk of blocking situations which can potentially arise during the legislative process because of a divergence of majorities between *Bundestag* (Federal Diet) and *Bundesrat* (Federal Council). The latter consists of at least three and at the most six members of every *Land* government and has the ability to block federal legislation whenever it affects the interests of the *Länder*. In that case, the Grundgesetz requires the *Bundesrat*'s consent, and the latter has an absolute veto which cannot be overridden by an equivalent vote of the *Bundestag* (Art. 77). In order to decrease the *Bundesrat*'s possibility to block federal lawmaking, the federalism reform has significantly reduced the percentage of laws requiring its consent. Besides that, it profoundly altered the rules concerning the division of legislative competences. While the previous so-called "framework" legislation was abolished, both the matters of the federation's exclusive (Arts. 71, 73) and concurrent (Art. 72) legislation were expanded.

The strong interconnection between federal and state level is at its greatest within the administration of justice. The courts of first and second instance (*Amtsgerichte*, *Landgerichte*, and *Oberlandesgerichte*) are state courts, whereas only the highest courts are established at the

federal level. Hence, there is a hierarchical division between the federal level and the *Länder*, and the federal courts do only control questions of legality. In addition to decentralization, there is also a high degree of specialization within the German court system. There are five different judicial branches – the ordinary, administrative, financial, labor, and social jurisdiction – with respective federal courts (*Bundesgerichte*) at their head which are scattered in different cities throughout Germany.

The German Civil Code and Private Law Society

The significance and importance of private law is emphasized by the ordoliberal concept of a Private Law Society (Böhm 1966), according to which the French Revolution marks the moment in which public order ceases to be the core governance mechanism in a hierarchically organized society, and consensus and private transactions become the main instruments for governing most parts of social life. However, since economic freedom, according to ordoliberal thinking, does only exist through order – it is an "ordered freedom" (Bonefeld 2012) – the new society, characterized by the equality of its members, party autonomy, freedom of contract, and freedom of competition, needs a strong state, called to assure the orderly conduct of self-interested entrepreneurs.

The German Civil Code of 1896 (Bürgerliches Gesetzbuch – BGB), which came into force on 1 January 1900 after 20 years of work, clearly reflects the core elements of a Private Law Society. The principle of freedom of contract, though it is nowhere expressly proclaimed, dominates the law of obligations and the idea that contracting parties are formally free and equal implicates that contracts thus formed must be adhered to in all cases. The typical citizen for the BGB is the *homo oeconomicus*, a person "who can be expected to have business experience and sound judgement, capable of succeeding in a bourgeois society with freedom of contract, freedom of establishment and freedom of competition" (Zweigert and Kötz 1994). Whenever the content of a contract was the result of a bargaining process, the former was considered to be fair – it was the contractual mechanism itself that was said to guarantee the correctness of its outcome (Schmidt-Rimpler 1941). Put succinctly, the bargaining process between the parties was seen as "the epitome of fairness" (Markesinis et al. 2006), and the parties were considered to be the best guarantors of their respective rights. Thus, the idea that individuals are free to engage in private transactions as they see fit only finds limits in the case of the violation of statutory prohibitions (§ 134), *bonos mores*, or in those cases in which one party has exploited the plight, inexperience, or lack of judgment of the other (§ 138), whereas the draftsmen expressly rejected the doctrine of *laesio enormis*, holding that the idea of a *iustum pretium* did not conform to the idea of freedom of contract.

Severe criticism was expressed by those who argued against the dominance of laissez-faire concepts in favor of a more solidaristic approach, calling for the limitation of freedom of contract, which was accused of favoring only the propertied classes while suppressing the socially weaker ones (Anton Menger). Nonetheless, at the end, only a few "drops of socialist oil" (Otto von Gierke) were added to the soulless individualism of the Code. Party autonomy should not be considerably restricted, neither by means of statutory nor judge-made law. In fact, the role that the few open-textured general clauses inserted into the BGB should once have played was a role the drafting fathers had not counted on.

The BGB is divided into five books, each of which is devoted to a different subject, with a complicated system of cross-referencing between them. This applies especially to the first book, the so-called General Part (*Allgemeiner Teil*) which contains certain basic institutions – such as the provisions concerning natural and legal persons (including the concepts of consumer (§ 13) and business (§ 14) which were introduced in 2000 by a statute implementing various EEC consumer directives), juristic entities and foundations, and general rules about legal acts (*Rechtsgeschäfte*) – that have an effect on all the other four books (*Law of Obligations*, *Law of Property*, *Family Law*, and *Law of Succession*).

In particular, the provisions regarding legal acts – an artificial concept which represents the leading topic of the *Allgemeiner Teil* – provide evidence of the BGB's abstract conceptual calculus. The fundamental component of a *Rechtsgeschäft* is the declaration of intent, and this is why it does not only include "normal" types of contract, such as sale or lease and the so-called real contracts necessary to create or transfer real rights over another's property, but also the contract of family law such as marriage, as well as unilateral acts such as that of making a will (Zweigert and Kötz 1994).

In its concepts and its language, the BGB is the child of the abstract conceptualism of the Pandectist school. It does not deal with particular cases in a concrete manner, but adopts an extremely abstract, logical, and accurate legal jargon which is admired for its rigor of thought and precision, but criticized because of its lack of comprehensibility, especially for the ordinary citizen.

The *BGB* During the Twentieth Century

The question immediately arising in this context is how a Code which was said to be more "the cadence of the Nineteenth than the upbeat to the Twentieth Century" (Radbruch), thus doing no more than prudently summing up the past rather than boldly anticipating the future (Zitelmann 1900), could have survived till this day, transiting the German Empire, the Weimar Republic, Nazi Germany, two world wars, and German reunification without being subjected to fundamental revision; a Code which was described and criticized as conceptualist, libertarian, unsocial and individualistic, and unaware of the changing economic and social climate in the late nineteenth century.

The few drops of socialist oil inserted into the BGB indeed soon proved to be insufficient, in particular in the field of labor law. Already during the Weimar Republic, it became clear that the existing regulation was inadequate and incomplete and that legislative reforms were necessary in order to better protect dependent workers by means of provisions concerning security of employment, worker participation, and minimum rates of pay (Zweigert and Kötz 1994). Thus, labor law for the most part developed outside the Code. But also in other areas of law, such as the law of housing, the need to take more account of the social needs of the time and to better protect the economically weaker parties in society led to a considerable increase of legislation outside the BGB; only since the 1960s parts of it have been gradually incorporated into the Code.

In certain cases, even the Code itself proved to be a useful instrument for coping with the changing social circumstances. In the aftermath of World War I, in a period of great instability, a sheet anchor to approach the arising great social and economic problems was represented by the general clauses inserted in the BGB. In particular § 242, which states that the performance of a contract has to be according to the requirements of good faith (*Treu und Glauben*), was envisaged by the judiciary as a possible loophole in order to afford the revaluation of debts, which proved necessary in consequence of the staggering hyperinflation of the early 1920s. Since currency laws explicitly established the nominalistic principle *Mark ist gleich Mark* (a paper Mark repaid during the inflation is equal to a gold Mark invested before inflation) and government rejected revaluation, it was left to the judiciary to find a solution for the constantly growing problem of the inflation effects on monetary relations. Although the legislator had never intended the principle of good faith to be used in this manner, in 1923, the *Reichsgericht* used it to request a revision of monetary obligations owed under contracts (RGZ 107, 78).

The significance and importance of the general clauses as a means which allows the courts to keep the Code up to date have also emerged in other areas. To cite just one example, in the context of standard terms of contract whose proliferation was the aftereffect of the Industrial Revolution at the end of the nineteenth century, the general clauses (especially § 242) turned out to be the decisive instrument in order to efficiently control the increasing number of cases dealing with standard terms. Many of the rules developed by the judiciary

in the course of time eventually turned into statutory form within the Act on General Conditions of Business of 1977 (Gesetz zur Regelung der Allgemeinen Geschäftsbedingungen – AGBG) whose provisions were partly inserted into the second book of the BGB in the course of the 2002 Act Modernizing the Law of Obligations (for details, see below).

However, the downside of open-textured principles with no clearly defined content became evident during the Nazi regime when the arbitrary misuse in particular of the principles of good faith and good morals, reinterpreted by the judges in order to direct the law in a way which served their nationalist ideology, led to the reassessment of a whole legal order by means of interpretation, resulting in the phenomenon of an *entartetes Recht* (Rüthers 1989), a "degenerated law," entirely based on the National Socialist *Weltanschauung*. Though the Nazis had initially warned against the general clauses, there was now a real proliferation of undetermined terms in a number of new measures and special laws in order to serve as "entry points" for their ideology. Even more, the *Führer* principle and the *völkische Rechtsidee* (the popular notion of the law) according to which law was no longer perceived in terms of the rights of the individual but as the rights of the people as determined by the state became supra-positive and preexisting principles able to limit the existing written law. Doctrinal rationality and legal certainty got completely lost (Stolleis 1998). The plans of the Academy of German Law to create a *Volksgesetzbuch*, a "people's Code," in order to replace the BGB, however, did not progress beyond the draft.

The BGB initially continued to be in force in both parts of Germany also after World War II. In the GDR, it was gradually replaced only some years after the construction of the Wall. On 1 April 1966, the Family Code led to the abrogation of the fourth book of the BGB; 10 years later, the BGB was entirely replaced by the Zivilgesetzbuch (ZGB) which entered into force on 1 January 1976 (and was replaced by the BGB in 1990). In the FRG, the entering into force of the Grundgesetz in 1949 has fundamentally inspired the evolution of German Private Law, through to its constitutionalization. In this context, again, the role of the general clauses was – and continues to be – a fundamental one. Today, the general clauses are commonly described as "gateways" (*Einfallstore)* for the Basic Law's value judgments (*Wertesystem*) in the realm of private law. In other words, the fundamental rights enshrined in the *Grundgesetz* do not have any direct effect on private law relationships, but it is the courts' duty to take them into account while interpreting the general clauses of the BGB in order to provide them with more concrete content. The idea of the fundamental rights' *mittelbare Drittwirkung* (indirect horizontal effect) was laid down for the first time in the famous Lüth Case of 1958 (7 BVerfGE 198), in which the BVerfG stated that the Basic Law contains an "objective order of values" (*objektive Werteordnung*) which must be looked upon as a fundamental constitutional decision affecting *all* areas of law, including private law.

The theory of the "radiating effect" (*Ausstrahlungswirkung*) of the basic rights in relations governed by private law has played a fundamental role in all subsequent court decisions on the subject and has given rise to a substantive body of case law developed by the courts. That it has also led to far-reaching consequences within the law of contract becomes particularly clear in the context of the famous Bürgschaft Case of 1993 in which the Federal Constitutional Court obliged civil courts to intervene on the basis of the general clauses of good morals and good faith – using the principles of party autonomy (Art. 2) and the social state (Art. 20) enshrined in the Constitution as interpretative guidelines – in a case in which the "structural imbalance of bargaining power" had led to an "exceptionally burdensome contract" for the weaker party (89 BVerfGE 214 (1993)).

The Europeanization of the BGB: Legal Developments in Private Law in the Twenty-First Century

The Act Modernizing the Law of Obligations *(Schuldrechtsmodernisierungsgesetz)* of 1 January 2002 is considered to be one of the most sweeping reforms of the BGB since it was enacted. The necessity to transpose several EC Directives,

among which the Consumer Sales Directive of 1999, was used as an occasion to realize the often called for fundamental reform of the law of obligations (the so-called *grobe Lösung*). This explains why most parts of the new provisions go back to and are largely based on reform proposals which had never progressed beyond the draft stage, prepared by a commission of leading academics and practicing lawyers in the early 1990s already. Because of its extremely short legislative process, the modernizing reform was heavily criticized by those who had argued in favor of a *kleine Lösung*, i.e., a legislation transforming only the Consumer Sales Directive without totally reforming the whole law of obligations. The Modernizing Act not only affected key elements of the general law of obligations and the contracts of the sale of goods, but also led to the integration into the BGB of many of the special statutes concerning consumer rights which over the years had developed outside the BGB and are the result of a transposition of EC Directives in this field. Since central parts of the new law of obligations have transposed European Directives, the reform is largely understood as a Europeanization of the BGB (Schulze and Schulte-Nölke 2001); the reform has thus also aimed to turn the BGB into a model in the context of the harmonization of European contract law. Having said that, the highly visible influence of the BGB within the framework of the Draft Common Frame of Reference (DCFR) is certainly also due to the leading role of German academics within the Joint Network on European Private Law entrusted with its elaboration.

The influence of EC/EU law has constantly grown over the last years, in particular in the field of consumer protection law which is increasingly inspired by the so-called information model. In German literature, the instrumentalization of the consumer as a means to promote the internal market on the basis of a model of procedural justice – more precisely of the idea that a sufficiently informed consumer is able to reach a completely rational decision – is said to have brought about a "dematerialization" and "re-formalization" of (national) consumer protection law (Micklitz 2012). The most recent example of this tendency is represented by the Consumer Rights Directive 2011/83/EU which was implemented in June 2013 by means of an amendment of the law of obligations.

Legal Education and Legal Professions: The German Construct of *Einheitsjurist*

The particularity and main characteristic of the German system of legal education is the fact that it is embedded in the model of a "uniform jurist," the so-called *Einheitsjurist*. The admission to all legal professions requires exactly the same formal qualification, so education is identical for everyone, regardless of the legal profession to be adopted in the future. Since federal legislation, i.e., the German statute of judges (*Deutsches Richtergesetz – DRiG*), ***only*** lays down general provisions and standards in the field of university education and practical training, each *Land* disposes of more detailed regulations. This has given rise to quite different rules – mainly related to procedural aspects of examination – in the 16 *Länder* (Keilmann 2006).

The legal education system is a two-phase one. It is a state-oriented and judge-centered bureaucratic model with traditionally great emphasis on the technical skills needed in the judiciary which has only been slightly reduced in recent times. It has its roots in the Prussian system of legal education and justice which already at the beginning of the eighteenth century was characterized by a state-controlled meritocratic selection for entry into the public administration on the basis of state examinations.

University studies in law end with an examination which today is called "first examination" (*erste Prüfung*) and no longer "first *state* exam" (*Erstes Staatsexamen*) because it is no longer entirely organized by the *Länder*'s Court of Appeals and the state offices for the Law Examinations (*Landesjustizprüfungsämter*). The modification brought about by the Act on the Reform of Legal Education of 2002 is an aggregate final exam mark, composed of an examination conducted by the law faculties in certain areas of specialization which students can choose (*universitäre Schwerpunktbereichsprüfung*) and a state exam which covers compulsory subjects (*staatliche Pflichtfachprüfung*). The former

G

represents 30, the latter 70 % of the final mark. Low average marks, high failure rates, and the assumption among students that university courses do not suitably prepare them for the exam have caused a run on the expensive service of private law teachers, the so-called *Repetitoren*. This dualism of abstract metatheory and practical legal training, which goes back to ideas diffused by Wilhelm von Humboldt at the beginning of the nineteenth century (Keilmann 2006), is heavily criticized but does not seem to draw to a close. Today, even the law faculties themselves offer special preparation classes in order to keep up with the private *Repetitoren* (Korioth 2006). The fact that privatization has affected the legal education system is also proved by the presence of the Bucerius Law School, the first and up to now the only private law school in Germany which was founded in 2000 and is located in Hamburg.

Legal studies are followed by a biennial practical training (*Referendariat* or *Vorbereitungsdienst*), organized and paid for by the *Land* in which it is undertaken. This period, during which the *Referendare* are instructed in various areas (civil courts, criminal courts or public prosecutor's office, public administration, and law firms) in order to obtain preparation for the main legal professions, ends with the "second state examination" (*zweite Staatsprüfung*). After successfully passing it – candidates only have two attempts, but failure rates are less than 15 % – the young lawyer is called *Assessor* or *Volljurist* (fully qualified jurist) and, having obtained "the qualification for the office of a judge" (§ 5 DRiG), he/she is theoretically qualified for any legal profession. That means, for example, that there is no additional bar exam, and to practice as a *Rechtsanwalt* (attorney), only a formal admission is required. However, access to a legal career is strongly determined by examination grades. Since for a career in the judiciary or civil service it is a precondition to achieve a top mark in the second state examination (the so-called *Prädikatsnote*) – a criteria which is fulfilled by fewer than 10 % of the aspirants – the overwhelming majority work as attorneys; only a small percentage has the option of becoming a judge (*Richter*), a public prosecutor (*Staatsanwalt*), a notary (*Notar*), or undertaking an academic career. As a response to this reality, the Reform of 2002 has extended the practical training during the *Referendariat* at a law firm from 3 to a minimum of 9 and a maximum of 13 months.

Attorneys who have practiced as lawyers for at least 3 years can specialize in a particular field of expertise in order to attain the title of *Fachanwalt*. In a few *Länder*, attorneys can also become the so-called *Anwaltsnotare* (lawyer notaries) who combine the profession of an attorney with that of a notary on the condition that they have practiced the forensic profession for 5 years. They are the second major type of notaries beside the so-called *Nurnotare* (fulltime notaries). Both groups have to pass a third exam, the so-called *notarielle Fachprüfung*. Due to the limited number of posts available, competition is hard, and the door to this profession is only open to the top law graduates.

As seen above, the judiciary follows a separate career path: it is immediately entered after the second state examination by those who have achieved excellent final marks, and only rarely does a practicing lawyer transfer to become a judge. The initial appointment to a local or a regional court will only be for a certain period, usually for 3 years, during which the candidate is a *Richter auf Probe*, i.e., his employment status is "on probation." A *Richter auf Lebenszeit*, a lifetime judge, can be promoted to higher regional courts on the condition that he has worked for at least 10 years in the courts of first instance. Whereas the initial appointments and this kind of promotion are made by the state's ministers of justice, appointments to the five federal courts involve the competent Federal Minister and a so-called *Richterwahlausschuss*, a committee for the selection of judges which consists of the competent *Land* ministers and an equal number of members elected by the *Bundestag*.

References

Böhm F (1966) Privatrechtsgesellschaft und Marktwirtschaft. ORDO 17:75–151

Bommarius C (2011) Lecture – Germany's Sozialstaat principle and the founding period. German Law J 12: 1879–1886. Available at http://www.germanlawjournal.

com/index.php?pageID=11&artID=1389. Bonefeld 2012
Bonefeld W (2012) Freedom and the strong state: on German ordoliberalism. New Polit Econ 17:633–656
Häberle P (1987) Die Menschenwürde als Grundlage der staatlichen Gemeinschaft. In: Isensee J, Kirchhof P (eds) Handbuch des Staatrechts, vol 1. C.F. Müller, Heidelberg, pp 815–861
Joerges C, Roedl F (2004) "Social market economy" as Europe's social model? European University Institute. Available at http://ssrn.com/abstract=635362
Keilmann A (2006) The Einheitsjurist – a German phenomenon. German Law J 7:293–312. Available at http://www.germanlawjournal.com/index.php?pageID=11&artID=712
Kielmansegg PG (1990) The basic law – response to the past or design for the future? In: Lehmann H, Ledford K (eds) Forty years of the Grundgesetz. German Historical Institute, Washington, DC, pp 5–18. Available at http://www.ghi-dc.org/publications/ghipubs/op/op01.pdf
Korioth S (2006) Legal education in Germany today. Wis Int Law J 24:85–107
Markesinis BS, Unberath H, Johnston A (2006) The German law of contract: a comparative treatise. Hart Publishing, Oxford/Portland
Micklitz HW (2012) Do consumers and businesses need a new architecture of consumer law? A thought-provoking impulse. European University Institute, Florence. Available at http://cadmus.eui.eu/bitstream/handle/1814/23275/LAW_2012_23_Rev.pdf?sequence=3
Müller-Armack A (1933) Staatsidee und Wirtschaftsordnung im neuen Reich. Juncker und Duennhaupt, Berlin
Müller-Armack A (1966) Wirtschaftsordnung und Wirtschaftspolitik. Rombach, Freiburg
Rüthers B (1989) *Entartetes Recht. Rechtslehren und Kronjuristen im Dritten Reich*. C.H. Beck, München
Scharpf F (2009) Föderalismusreform: kein Ausweg aus der Politikverflechtungsfalle? Campus-Verlag, Frankfurt/New York
Schmidt-Rimpler W (1941) Grundfragen einer Erneuerung des Vertragsrechts. Archiv für die civilistische Praxis 147:130–197
Schulze R, Schulte-Nölke H (eds) (2001) Die Schuldrechtsreform vor dem Hintergrund des Gemeinschaftsrechts. Mohr Siebeck, Tübingen
Somma A (2013) Private law as biopolitics: ordoliberalism, social market economy, and the public dimension of contract. Law Contemp Probl 76:105–116. Available at http://scholarship.law.duke.edu/lcp/vol76/iss2
Stolleis M (1998) The law under the swastika. The University of Chicago Press, Chicago
Stolleis M (2002) Geschichte des öffentlichen Rechts in Deutschland: Weimarer Republik und Nationalsozialismus. C.H. Beck, München
Zacher HF (1987) Das soziale Staatsziel. In: Isensee J, Kirchhof P (eds) Handbuch des Staatsrechts der Bundesrepublik Deutschland, vol 1. C.F. Müller, Heidelberg, pp 1045–1112
Zitelmann E (1900) Zur Begrüβung des neuen Gesetzbuches. Deutsche Juristen-Zeitung 5:2–6
Zweigert K, Kötz H (1994) An introduction to comparative law. Clarendon, Oxford

Gibbard-Satterthwaite Theorem

Pierre Bernhard[1] and Marc Deschamps[2]
[1]Biocore team, Université Côte d'Azur-INRIA, Sophia Antipolis Cedex, France
[2]CRESE EA3190, Université Bourgogne Franche-Comté, Besançon, France

G

Definition

One seminal question in social choice theory was: is it possible to find a social choice function such that each agent is always better off when telling the truth concerning his preferences no matter what the others report? In other words, can we find a strategy-proof voting rule? With at least three alternatives and two voters, the answer is clearly no under a very general framework, as was proved independently by Allan Gibbard and Mark Satterthwaite. Since then, the Gibbard-Satterthwaite theorem is at the core of social choice theory, game theory, and mechanism design.

Introduction

Since K. Arrow's 1951 analysis, which marks the revival of the theory of social choice, economists investigate from an axiomatic point of view the aggregation of individual preferences in order to obtain a social welfare function (i.e., a complete and transitive ranking based on the individual preferences) or a social choice function (i.e., one alternative from the individual preferences). Such questions concern huge domains of human beings, for example, the family's choice of the walls color in the living room, the choice of the Palme d'Or winner by the jury of the Cannes International Film Festival, or the vote for the President of the European Commission. Thus, in

addition to his (im)possibility theorem, K. Arrow has initiated a new domain in economic analysis.

Since then, the question of the strategic behavior of individuals during an election was raised again on several levels. Indeed for a very long time, since we find, for example, a letter from Pline the Younger in Roman Antiquity, the question of manipulation had attracted attention. Whether playing on who will be a voter, who can be a candidate, the choice of the voting rule, abstention, beliefs about preferences or preferences expressed during the vote, the idea that at least one individual can do a manipulation to his advantage has interested and concerned many thinkers.

Concerning this last form of manipulation, it was historically mentioned at least seven times before the Gibbard-Satterthwaite theorem (an expression first used by (Schmeidler and Sonnenschein 1978)). First, a story probably apocryphal tells that Borda responded to one of his critics who pointed out that his method was manipulable that it was intended for use by honest people. There is also a reference to this question in Ch. Dodgson (alias Lewis Carroll), who said in a specific voting system, "This voting principle makes an election more of a skill game than real test of voters' wishes" (see Black (1958)). In the modern period, Black (1948) discusses the link between unimodal preferences and strategic votes, while (Arrow 1951, p. 7) explicitly states that he will not deal with this issue even though he later returns to it in a footnote [footnote 8, pp. 80–81]. Arrow's general analysis, however, will lead (Vickrey 1960, pp. 517–519) to conjecture that immunity to strategic manipulation is logically equivalent to the association of the axiom of independence of irrelevant alternatives and that of positive association. Finally, it will be R. Farquharson in his 1958 doctoral dissertation, published in 1969, (Farquharson 1969), who will introduce the distinction between "sophisticated strategy" and "sincere strategy," and then in an article with M. Dummett in 1961 when they make the conjecture: "It seems unlikely that there is any voting procedure in which it can never be advantageous for any voter to vote 'strategically,' i.e., non sincerely" (Dummett and Farquharson 1961, p. 34). In an interview given in 2006 to R. Fara and M. Salles, M. Dummett confesses that he felt at this time that proving this conjecture would be extremely difficult and that is why they did not try the demonstration. We refer the reader to (Barberà 2010) for more details on this historical part.

Thus, for more than 20 years after (Arrow 1951), all scholars of social choice theory seem to have been convinced that the question of the manipulation of preferences by an individual (i.e., the fact that he does not express his true preferences in order to lead to a social choice that satisfies him better than would have been obtained if he had been honest) was an important question, easy to explain, but very difficult to prove.

As an illustration of this question of the manipulation of preferences by an individual, we can give the following example (from (Feldman 1979, p. 459)):

Imagine a committee made up of 21 members each having one vote and whose actual preferences presented in descending order can be broken down into three groups.

Type 1	Type 2	Type 3
A	B	C
B	C	B
C	A	A
10 voters	9 voters	2 voters

In a majority election where voters indicate their real preferences, A gets 10 votes, B gets 9 votes, and C gets 2 votes; this leads to the election of A. However, if anticipating this result, the voters in group 3 manipulate their preferences and vote for B, this will lead to 10 voters for A and 11 voters for B and so B will be elected. This strategic choice of the voters of the group 3 allows them to obtain a more favorable result than they would have obtained if they had voted sincerely. ∎

Therefore, a question immediately comes to mind: is it possible to imagine a social choice function such that each individual, regardless of the others' choices, is always better off by expressing his true preferences? In other words, is there a social choice function such that always telling the truth is a strictly dominant strategy for

each individual? The answer to this question was independently given by the philosopher Allan Gibbard (1973) and the economist Mark Satterthwaite (1975) (work summarizing part of his doctoral dissertation defended in 1973) and it is unfortunately negative. What we call since the Gibbard-Satterthwaite theorem can be broadly stated as follows: *it is not possible to find a social choice function that is both nonmanipulable and nontrivial*.

This last term deserves an explanation. By a trivial function we mean one of the following solutions (or a solution equivalent to it): 1/ a dictatorship (i.e., all the power of decision resides with a single individual and he has no indifferent choice), 2/ the permanent choice of an alternative whatever the preferences expressed by the individuals are (which could, for example, correspond to a tradition which would be imposed in all circumstances), 3/ the perfect unanimity of all the voters (which in fact means having perfect clones and therefore no diversity in the preferences), or 4/ the majority between two alternatives only, whatever the number of other alternatives existing and the votes that are expressed for them. Thus, if all individuals know their preferences and those of others and that there are at least three possible alternatives and at least two voters, there is no social choice function that implies that each individual always has an interest in expressing his or her real preferences. We thus find here the conditions of Arrow's theorem and the same conclusion since when there are only two alternatives, majority voting is at the same time a nondictatorial and nonmanipulable social choice function.

Since then, the Gibbard-Satterthwaite theorem has been considered, along with Arrow's theorem, as one of the two most famous results of social choice theory and has led to a very large literature in this field. It also plays a crucial role in public economics and in the theory of incentive mechanisms, which can be broadly seen as a social engineering approach of finding the rules to achieve a specific outcome from agents interacting strategically and having private information (see Börger (2015)). Indeed, because of this theorem, the incentives of individuals must be considered as relevant constraints in the design of any mechanism.

Gibbard-Satterthwaite Theorem

Many proofs of this theorem have been proposed and it is possible to consider that they take one of the following four paths: 1/ that used by A. Gibbard and which uses Arrow's theorem, 2/ that used by M. Satterthwaite thanks to a combinatorial argument and recurrences on the number of individuals and alternatives, 3/ that considering this theorem as the consequence of the fact that nonmanipulability requires strong monotony (see Moulin (1988)), and 4/ that developed by S. Barberá and his coauthors using the concept of pivotal agents. For a first presentation of the question of manipulation, we refer to Feldman (1979) and for a review of this literature we refer to Sprumont (1995) and Barberà (2010).

Following Gibbard's approach, we will prove the Gibbard-Satterthwaite theorem as a corollary of Arrow's (im)possibility theorem. The presentation we retain will distinguish the formal setup, the links between the two theorems, the proof of the Gibbard-Satterthwaite theorem, and that Arrow's theorem in turn can be seen as a corollary if we directly prove the Gibbard-Satterthwaite theorem.

Formal Set-up

Let $\mathcal{A} = \{a, b, \ldots\}$ be a set of three or more *alternatives*. Let P be the set of linear orders over $\mathcal{A}$, and $\mathcal{P} = P^n$. An element $\Pi = (\mathrm{P}_1, \mathrm{P}_2, \ldots, \mathrm{P}_n) \in \mathcal{P}$ is called a *profile*, and the P_i the *individual preferences*. The order in P_i is denoted $a \succ_i b$ for "player i prefers a to b." We further define

Definition 1
The domination set *of* $a \in \mathcal{A}$ *in* P_i

$$\mathcal{D}(a, P_i) = \{x \in \mathcal{A} |\, a \succ_i x\}$$

Moreover, for $\Pi = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$*,* $\mathcal{D}(a, \Pi)$ *stands for the product set of the* $\mathcal{D}(a, \mathrm{P}_i)$*. Thus,*

let $\Pi' = (P', \ldots, P'_n)$ *be another profile, the notation* $\mathcal{D}(a, \Pi') \supseteq \mathcal{D}(a, \Pi)$ *means:* $\forall i \in \{1, \ldots, n\}$, $\mathcal{D}(a, P'i) \supseteq \mathcal{D}(a, P_i)$.

Definition 2

- *A* social welfare function *(or SWF) is an application* $f: \mathcal{P} \to P$.
- *A* social choice function *(or SCF) is an application* $F: \mathcal{P} \to A$.

The order relation in $f(\Pi)$ will be denoted $a > b$ or, if needed $a >_{f(\Pi)} b$. We need to define the following properties of SWF or SCF:

Definition 3

- *The SWF f is* Pareto efficient *(or* satisfies the unanimity rule*) if*

$$[\forall i \leq n, \ a \succ_i b] \Rightarrow a > b.$$

- *The SCF F is* Pareto efficient*, (or* satisfies the unanimity rule*) if whenever an alternative* $a \in \mathcal{A}$ *is the most preferred alternative in all individual preferences, it results that* $F(\Pi) = a$.
- *The SWF f is* independent of irrelevant alternatives *(IIA) if, for any a and b in* $\mathcal{A}$*, the relative ranking of a and b in* $f(\Pi)$ *only depends on their relative rankings in the* P_i*, irrespective of the rankings of other alternatives.*
- *The SCF F is* monotonic *if, given two profiles* Π *and* Π',

$$[F(\Pi) = a \text{ and } \mathcal{D}(a,\Pi') \supseteq \mathcal{D}(a,\Pi)] \Rightarrow F(\Pi') = a.$$

- *The SCF F is* strategy-proof *if, when* Π *and* Π' *differ only in changing* P_i *into* P'_i*, it results that either* $F(\Pi) = F(\Pi')$ *or* $F(\Pi) \succ_i F(\Pi')$.
- *The SWF f is called* dictatorial *if there exists a player k such that,* $\forall \Pi \in \mathcal{P}, f(\Pi) = P_k$. *(The social preferences are always player k' preferences.)*
- *The SCF F is called* dictatorial *if there exists a player k such that,* $\forall \Pi \in \mathcal{P}$, $\forall x \neq F(\Pi)$, $F(\Pi) \succ_k x$. *(F (Π) is player k's most preferred alternative.)*

Arrow and Gibbard-Satterthwaite Theorems

Theorem 1 (Arrow) *If a SWF is Pareto efficient and IIA, then it is dictatorial.*

Theorem 2 (Gibbard-Satterthwaite) *If a SCF is strategy-proof and onto (i.e., its range is all of* $\mathcal{A}$*:* $\forall a \in \mathcal{A}, \exists \Pi \in \mathcal{P}$ *such that* $F(\Pi) = a$*), it is dictatorial.*

Lemma 1 (Muller and Satterthwaite) *If a SCF is strategy-proof and onto, it is monotonic and Pareto efficient.*

Proof Let Π and Π' be two profiles, $F(\Pi) = a$, $\mathcal{D}(a, \Pi') \supseteq D(a, \Pi)$, but assume that $F(\Pi') \neq a$. Make the change from Π to Π' one player at a time in numeric order. Denote by Π_k the profile obtained after changing P_k to P'_k. (And $\Pi_0 = \Pi$). At some point, we have that $F(\Pi_{i-1}) = a \neq b = F(\Pi_i)$. By strategy-proofness, it follows that $a \succ_i b$ in P_i while $b \succ_i a$ in P'. But by hypothesis, if $a \succ_i b$ in P_i it is a fortiori true in Π', a contradiction. Therefore, the SCF is monotonic.

Because F is assumed to be an onto function, for any given $a \in \mathcal{A}$, there exists a profile Π such that $F(\Pi) = a$. Build Π' by moving a at the top of the preferences of all players. By monotonicity, it still holds that $F(\Pi') = a$. Now, get Π'' by shuffling at will the preferences of all players *below a*, leaving a at their top position. Because of the definition of monotonicity, and specifically its IIA-like character, it still holds that $F(\Pi'') = a$. Therefore, the SCF is Pareto efficient. **I**

Proof of the Gibbard-Satterthwaite Theorem

This note is devoted to the proof of the Gibbard-Satterthwaite theorem viewed as a corollary of Arrow's theorem. We assume therefore that the latter is known. Given the above Lemma 1, we need to prove

Lemma 2 *If a SCF is Pareto efficient and monotonic, it is dictatorial.*

Or method of proof will be as follows: we assume that a SCF F is known which is Pareto efficient and monotonic. From it we construct a SWF f which we show to be Pareto efficient and IIA, therefore dictatorial, which will imply that F is dictatorial.

Let F be a SCF. Given a profile Π, our social ordering $f(\Pi)$ is built via the following algorithm:

Algorithm

- Let $\Pi_1 = \Pi$.
- For i ranging from 1 to n, do:
  - define $a_i = F(\Pi_i)$,
  - define Π_{i+1} by moving a_i at the bottom of all individual preferences in Π_i.
- Define $f(\Pi)$ as: for all $i \in \{1, \ldots, n-1\}$, $a_i > a_{i+1}$.

Proposition 1 *If the SWF F is Pareto efficient and monotonic, the above algorithm produces a linear order on $\mathcal{A}$.*

Proof What we need to prove is that all elements of A will be ranked, i.e., that for all $i \in \{1, \ldots, n\}$, and for all $j < i$, $a_i \neq a_j$.

Let first $i < n$. Therefore, not all alternatives have been numbered as one of the a_k. Define Π_i' as follows: Take an alternative b which has not yet been ranked. Raise it at the top of all individual preferences. This does not change the domination sets: $\mathcal{D}(a_j, \Pi_i') = \mathcal{D}(a_j, \Pi_i)$ since in Π_i, all the alternatives that have already been selected in the algorithm, including a_j (remember that $j < i$), are stacked at the bottom of all individual preferences. Therefore, by monotonicity, if $F(\Pi_i) = a_j$, it also holds that $F(\Pi_i') = a_j$. But by Pareto efficiency, $F(\Pi_i') = b$, a contradiction.

Finally, in Π_n, $n-1$ different alternatives have been placed at the bottom of all individual preferences; thus, the same and last alternative is alone at the top of all and is therefore selected by F by Pareto efficiency.

The following propositions end our proof:

Proposition 2 *If the SCF F is Pareto efficient and monotonic, the SWF defined by the algorithm is*
1. *Pareto efficient*
2. *Independent of irrelevant alternatives (IIA), and therefore dictatorial by Arrow's theorem*

Proof

1. Let $a, b \in \mathcal{A}$, and assume that in Π, $\forall i \in \{1, \ldots, n\}$, $a \succ_i b$. Assume that at some step of our algorithm, $F(\Pi_i) = b$, but a has not yet been chosen. In Π_i, raise alternative a at the top of all individual preferences. This does not change the domination sets of b in any individual preferences. Therefore, F should still select b. But by Pareto optimality, it should select a. A contradiction.
2. Let $i < j$ and therefore by our algorithm, $a_i >_{f(\Pi)} a_j$. In Π, move another alternative b in some individual preferences, call the new profile Π'. We claim that in the order created by our algorithm applied to Π', it still holds that $a_i > a_j$.

Assume that $F(\Pi') = b$. Then at step 2, b is brought at the bottom of all individual preferences, and, as compared to profile Π, the domination sets of a_1 have been enlarged or kept unchanged for all individual preferences. Therefore, $F(\Pi_2') = a_1$. If $b \neq a_2$, at step 2, for the same reason a_2 will be selected, and so forth until the end of the algorithm. Therefore, we will still select a_i before a_j.

Assume $F(\Pi') = c \neq b$, and assume $c \neq a_1$. This is possible only if in one individual preferences at least, b has been moved from below a_1 to above it. In these individual preferences only, bring b back below a_1. Call this profile Π''. $\mathcal{D}(a_1, \Pi'') \supseteq \mathcal{D}(a_1, \Pi)$. Therefore $F(\Pi'') = a_1$. But in going from Π' to Π'', b has only been moved down in some individual preferences. Therefore $\mathcal{D}(c, \Pi'') \supseteq \mathcal{D}(c, \Pi')$. Therefore, $F(\Pi'') = c$. Hence, $a_1 = c$ contrary to the hypothesis. Hence, $F(\Pi') = a_1$.

Repeat this argument at each step before b is selected, and once this happens, repeat the

previous argument. It follows that, except for b, all other alternatives are chosen in the same order as for Π. ∎

It follows that, by Arrow's theorem, f is dictatorial. There is a player k such that for all Π, $f(\Pi) = P_k$, and in particular $F(\Pi)$ is player k's preferred alternative. And this proves Lemma 2, and consequently the Gibbard-Satterthwaite theorem. ∎

Complement: Arrow's Theorem as a Corollary of Gibbard Satterthwaite

It can also be shown that if the Gibbard-Satterthwaite theorem has been proved directly, Arrow's theorem is an easy corollary, thus establishing the equivalence between the two results. The process is symmetrical from the above one, deriving a SCF from a Pareto efficient and IIA SWF by simply picking the top alternative in the social preferences and showing that it is onto and strategy-proof.

We might also mention that Lemma 1 has an easy reciprocal.

Conclusion

Since its demonstration, the Gibbard-Satterthwaite theorem has generated a huge literature on the question of the manipulation of preferences, distinguishing, in particular, the cases where the social choice function is manipulable by an individual from the case where it is manipulable by a coalition of individuals. It has also been extended to take into account set-valued (Duggan and Schwartz 2000) and nondeterministic choice functions (Gibbard 1977), that is, those where the result depends on the votes of individuals but also by chance. (Often both set-valued and nondeterministic.) Our framework requires that social choice function is onto. It has been proven since that the results hold as soon as the social choice function has at least three alternatives in its range.

In our opinion, three main ways of circumventing this theorem have been followed. The first is to restrict the preference domain (with functions defined on restricted sets of preference profiles, it is possible to find social choice functions that are both nondictatorial and nonmanipulable, e.g., unimodal preferences *à la* Black). The second is to change the goal. Indeed, the framework in which the Gibbard-Satterthwaite theorem is situated is very strong since it seeks a social choice function for which telling the truth is a dominant strategy for each individual. The path followed by implementation theory is to simply ask that it be a Nash strategy, or a perfect subgame strategy, or a Bayesian Nash strategy, depending on the informational context of the individuals. Finally, more recently, a third way explains that this problem is real but may not be very important empirically. Indeed, besides the integrity, ignorance or stupidity of individuals that can prevent them from performing manipulations, the fact that a social choice function is manipulable does not imply that it will be manipulated. And since Bartholdi et al. (1989), economists consider that it may be empirically impossible for individuals to decide how to manipulate even when they have all the information to do so, as the problem may be NP-hard.

Cross-References

- ▶ Electoral Systems
- ▶ Heterarchy
- ▶ Liberty
- ▶ Political Competition
- ▶ Political Economy
- ▶ Politicians
- ▶ Public Goods
- ▶ Public Interest
- ▶ Rationality
- ▶ Simple Majority

References

Arrow KJ (1951) Social choice and individual values. Wiley, New York

Barberà S (2010) Strategy-proof social choice. Barcelona economic working paper 420. University of Barcelona, Spain

Bartholdi J, Tovey C, Trick M (1989) The computational difficulty of manipulating an election. Soc Choice Welf 6:227–241

Black D (1948) On the rationale of group decision making. J Polit Econ 56:23–34
Black D (1958) The theory of committees and elections. Cambridge University Press, Cambridge, UK
Börger T (2015) An introduction to the theory of mechanism design. Oxford University Press, New York
Duggan J, Schwartz T (2000) Strategic manipulability without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized. Soc Choice Welfare 17:86–93
Dummett M, Farquharson R (1961) Stability in voting. Econometrica 29:33–44
Farquharson R (1969) Theory of voting. Yale university Press, New Haven
Feldman AM (1979) Manipulating voting procedures. Econ Inq 17:452–474
Gibbard A (1973) Manipulation of voting schemes: a general result. Econometrica 41:587–601
Gibbard A (1977) Manipulation of schemes that mix voting with chance. Econometrica 45:665–682
Moulin H (1988) Axioms of cooperative decision making, Economic Society monographs. Cambridge University Press, Cambridge, MA
Satterthwaite M (1975) Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. J Econ Theory 10:187–217
Schmeidler D, Sonnenschein H (1978) Two proofs of the gibbard-satterthwaite theorem on the possibility of a strategy-proof social choice function. In: Gottinger HW, Leinfellner W (eds) Decision theory and social ethics, Issues in Social Choic. Reidel Publishing Company, pp 227–234. Published version of a 1974 working paper. Dordrecht:Holland/Boston: USA London: England
Sprumont Y (1995) Strategy-proof collective choice in economic and political environments. Can J Econ 28:68–107
Vickrey W (1960) Utility, strategy and social decision rules. Q J Econ 74:507–535

Global Warming

Alfred Endres
FernUniversität Hagen, Hagen, Germany

Abstract

This essay analyses global warming from the perspective of environmental economic theory: National activities to curb greenhouse gas emissions are public goods. It is well known that the expectations to arrive at a satisfactory provision of these kinds of goods by uncoordinated decision-making are low. Nations try to overcome this dilemma by international treaty-making. However, economic analysis suggests that the genesis and persistence of international climate agreements are challenged by problems of national rationality and stability as well as by the leakage effect. On the other hand, economic analysis reveals some reasons for a more optimistic view. Among those are incentives to develop and introduce green technologies and incentives to adapt to changes in the world climate.

G

Synonyms

Greenhouse effect

Introduction

Human economic activity is directed to produce commodities and services which are valued by people. Many of these goods are sold and bought in markets where consumers pay a market price to reimburse the production costs to the supplying firms and contribute to the profits of these firms. Jointly with these intended products of the economic activity, unwanted side products are generated, like waste and emissions into environmental media (e.g., air and water). The generation of these side products can be steered in terms of quantity and quality, but it is unavoidable in principle.

There are millions of different types of substances which are emitted into the environment as by-products of the productions process. A certain subset of these emissions constitutes the group of *greenhouse gases (GHG)*. Among this type of gases are carbon dioxide (CO_2), nitrous oxide (N_2O), chlorofluorocarbons (CFCs), methane (CH_4), and ozone (O_3). There is a broad consensus in the natural sciences that the emission of these greenhouse gases changes the atmospheric conditions around the globe to such an extent that the world climate is adversely affected. Particularly, increases in the stock of greenhouse gases in the atmosphere contribute to an increase of the average temperature on earth (*global warming*). This

is a phenomenon of fundamental concern in the society and for policy makers. It is also an important issue of research undertaken and policy advice given by economists. Climate change affects the living conditions of the people directly, and it also changes the conditions under which many commodities and services are produced (e.g., in agriculture and tourism).

From the perspective of sciences, climate change is an issue which has so many dimensions intertwined in a very complex way that it can only be dealt with using an interdisciplinary approach. Of course, in this interdisciplinary process of communication, collaborators from the social sciences (like economics) cannot really assess the quality of the research in the natural sciences. So economists usually take for granted what the mainstream of natural science research takes to be true. This is particularly so with regard to the quantitive effect that certain greenhouse gases have on the world climate and in terms of the effect that certain changes in the climate have on the living conditions on earth (e.g., sea level rise and increases in extremely adverse weather conditions).

There is a broad consensus in the world that global warming has anthropogenic origins and is extremely dangerous to human kind. Even though not each consequence of higher average temperature in the world is detrimental, the aggregate effect appears to be a severe threat to the living conditions on earth. (A standard reference in this context is the *Stern Review*, Stern (2007), see also Stern (2015))

The problem is very well known among the political leaders of the world and to the people, at least in the economically developed countries. (The beginning of the international efforts to protect the climate was marked by the *Earth Summit* in Rio de Janeiro in 1992.) On the other hand, there is no progress in terms of a reduction in the globally aggregated quantity of GHG emissions. In the public media, “uninterested and incompetent politicians” are often held responsible for this lack of progress. However, economic analysis shows that the climate problem is very hard to solve even if we assume that the involved governmental decision-makers are competent and seek to serve the interest of the people they represent. This is due to an adverse incentive structure that impedes an international consensus on an effective climate change policy.

An Economic Perspective of Global Warming

Greenhouse gases are to be distinguished from most other pollutants in that they do not have a regionally specific dose–response profile. Instead, they are global in the sense that after their emission they disperse homogeneously around the atmosphere of the earth. Particularly, the regional structure of these pollutants’ generation is irrelevant for their effect on the global climate. Changes in the world climate are induced by the aggregate effect of GHG emissions and, particularly, by the stock of greenhouse gases in the atmosphere that is accumulated by the emission flow generated in each consecutive period of time. Thereby, global pollutants have properties that come very close to the economics textbook stylization of a “pure public good (bad),” which is an important issue in *microeconomic theory*, *public finance,* and *environmental economics*. The problems that arise when governments strive to agree upon curbing GHG emissions to a tolerable level can thereby be explained using insights from the economic theory of public goods.

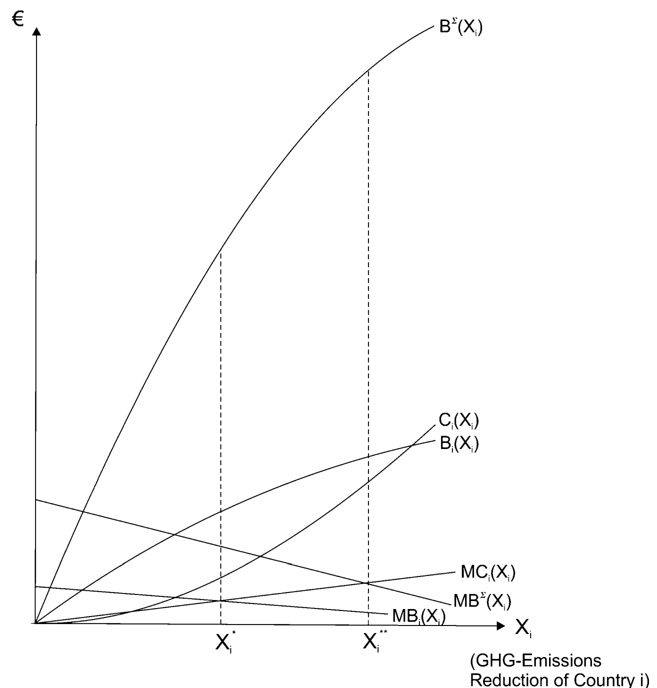
Concerning greenhouse gases, each individual country is in a double role: On the one hand, the country is the generator of the problem emitting certain quantities of these gases. On the other hand, the country is the victim of its own activity suffering from the detrimental effect this activity has on the climate. However, since it is not a country’s specific climate which is under consideration here but the world climate, the country is not the only victim of its own activity. All other countries are also victimized by the pollution of the individual country under consideration. In terms of economic theory, each country generates an *internal effect* as well as an *external effect* by its GHG emissions. The problem arising here is that a government striving to maximize the welfare of the citizens that it represents designs a GHG

policy that takes the internal effect into account and ignores the external one.

The welfare effects of a country's GHG policy are two-fold. On the one hand, the national economy has (the citizens have) to bear the costs of the policy. This reduces national welfare. On the other hand, the country benefits from its own policy in that the GHG-reducing measures are advantageous in terms of the world climate. Therefore, the climate change damages the country under consideration suffers from are reduced. The national welfare effect of the national climate policy is maximized when this policy is designed to maximize the difference between the national benefits (reduction in national climate change damages) and national climate change costs. The quantity of GHG reduction for which the aforementioned difference is maximized is called the *equilibrium quantity* of emission reduction for the country under consideration. It is fundamental to note that this nationally optimally designed climate policy is drastically different from the design of a national climate policy which is optimal from the point of view of the world as a whole. From this global point of view, in order to characterize optimal national climate policy, the cost of this policy has to be compared with the benefits of this policy aggregated across all the countries of the world (instead of the country under consideration only). The quantity of country i's emission reduction which maximizes the difference between aggregate benefits and costs to country i is called the *globally optimal quantity* of i's emission reduction.

In economics, it is a standard exercise to clarify and specify these thoughts using formal mathematical analysis or graphical visualization. In what follows, we turn to the latter Fig. 1.

In Fig. 1, the term B_i denotes the benefit, B, of GHG emission reduction in some country, i. X_i denotes the quantity of this reduction in the aforementioned country. The function $B_i(X_i)$ denotes these benefits as they depend upon the quantity of the pollutants reduced. It is assumed in this figure that this function is known to the decision-makers. In reality (and in more elaborate theoretical expositions), it is to be acknowledged that the extent of damages and the manner in which they systematically depend on the extent of emissions reduction is highly uncertain. There is an extensive literature dealing with this issue, which is ignored here. In the present context, the figure is



Global Warming, Fig. 1 The Benefits and Costs of GHG abatment: The National and the Global

used to illustrate the divergency between globally optimal emission reductions and the emission reductions chosen by individual countries striving to maximize the welfare of their respective citizens. In this context, it is sufficient to interpret the curves in Fig. 1 to be the best estimate of the relationships dealt with here that is available to the political decision-makers. Of course, a benefit curve like the one that has been drawn for country i can be drawn for any other country. This is not done in the figure. However, a curve $B^{\Sigma}(X_i)$ is drawn representing the vertical aggregation of the individual benefit curves of all countries. This aggregate curve shows how the sum of worldwide benefits from pollution reduction depends upon emissions reduction of country i. Additionally, Fig. 1 presents the curve $C_i(X_i)$ representing how the costs of emission reduction in country i depend upon the quantity of emissions that this country reduces. Comparing these curves, it is obvious that the globally optimal level of GHG reductions in country i is at X_i^{**}, the quantity where the difference between the aggregate benefit curve and the curve representing the abatement cost of country i is maximized. The quantity of emission reductions maximizing the welfare in country i (the "equilibrium quantity" for country i), however, is at X_i^*. In addition to these curves representing total benefits and costs, the figure shows the curves $MB_i(X_i)$, $MB^{\Sigma}(X_i)$, $MC_i(X_i)$. These curves are "partner curves" to the aforementioned curves in that they represent the marginal country-specific benefits, marginal aggregate benefits, and marginal country-specific costs instead of the total benefits and costs. These curves represent the country-specific benefits, aggregate benefits, and country-specific costs, respectively, for small increases of the quantity of emissions reduced. Technically speaking, these marginal curves are the first-order derivatives of the total curves. The equilibrium emission reduction quantity, X_i^*, for country i is defined by the intersection of the country's specific marginal benefit curve with the country's specific marginal cost curve. For any quantity lower than X_i^*, the marginal benefit of the last unit's reduction is higher than the cost of this last unit. Therefore, national welfare could be increased by reducing an additional unit. It follows that the starting point emission reduction level in this little thought experiment cannot be the equilibrium one. Analogously, for any level of pollution abatement beyond X_i^*, it can be argued that national welfare increases when pollution abatement is attenuated. By analogous reasoning, it can be shown that the globally optimal quantity of emission reduction, X_i^{**}, is defined by the intersection of the marginal aggregate benefit curve, $MB^{\Sigma}(X_i)$, and the marginal country-specific cost curve, $MC_i(X_i)$. In economic analysis (as applied to environmental problems and very many other issues), marginal analysis is a very common and powerful tool. It will be repeatedly used in a subsequent exposition.

International Environmental Agreements – Painful Genesis and Fragile Architecture

In the preceding section, the decisions of national governments have been stylized regarding the quantity of GHG reductions. It has been shown that a national GHG policy striving to maximize national welfare is unable to meet the global requirements of effective GHG reduction.

Even though we obviously do not have a "world government," it is instructive to imagine what a government like that might do. The result of this reasoning may be used as a benchmark which serves for the results of real world decision processes. To build up this benchmark, it is usually assumed in economics that this world government is well informed and benevolent. Applying the established economic welfare criteria, the world government would choose a situation with two essential properties. The first property is that GHG emissions would be at their globally optimal level as characterized in the preceding section. The second property relates to the question, how much each individual country would have to contribute to the globally optimal aggregate emission reduction. Thriving to keep the worldwide total costs of GHG abatement at their minimum, the world government would put a burden to an individual country which would be

the higher, the lower the abatement cost of this country is. More precisely, the cost-effective distribution of country-specific emission reduction loads is characterized by the marginal abatement costs to be equal across all countries. This "equimarginal condition" is very often used in economics (within the environmental context and beyond) and is explained more elaborately, e.g., in Endres (2011), pp. 216–220. If the countries of the world are aware of what has just been said and note that this solution cannot be achieved by country-specific "autistic" welfare maximization, they might try to get there by *voluntary agreement*.

The fundamental challenge to the design of international agreements to fight global warming is that they must be *self-enforcing* since international law is too weak to implement enforcement. This means that the international contract must be designed in a manner that it is attractive for each country to join in, and having joined, to keep the commitments. These two properties are called the *individual rationality* and the *stability* of an international treaty.

Very unfortunately, it does not follow from the global optimality property of an international GHG treaty that it is also individually rational and stable. To the contrary, for many countries there is an incentive to "cheat" upon the commitments they have signed in an international agreement by taking a *free ride* on the GHG abatement efforts of the other partaking countries (lacking stability). Moreover, some countries might have an incentive not to sign the treaty at all, even though this treaty is optimal from the global point of view (lacking individual rationality). This is particularly so for countries with low marginal abatement costs which do not suffer very much from global warming (or may even benefit from it).

There are many attempts in the literature to fill these rather theoretical considerations with empirical content. A prominent example is a study by Finus/van Ierland/Dellink (2006). (There is a textbook-style interpretation of the aforementioned original research in Endres (2011).) To investigate the incentive compatibility of a globally optimal GHG agreement, the authors of this study proceed as follows. They divide the world into 12 "regions" and use data from the empirical literature to estimate the GHG abatement benefit and cost functions for each of those regions. These 12 regions (countries or groups of countries, respectively) are USA, Japan, European Union, other OECD countries, Central and Eastern European countries, former Soviet Union, energy-exporting countries, China, India, dynamic Asian economies, Brazil, and "rest of the world". The result in terms of individual rationality is that three regions would deteriorate their welfare position when the situation would change from the pre-treaty situation (with each country maximizing its own welfare "autistically") to the globally optimal contract. These regions are the "other OECD countries," the "Central and Eastern European countries," and the countries of the former Soviet Union. For these, it is not individually rational to join a coalition that agrees to realize the globally optimal global warming policy. The results of the stability test are even less encouraging: With the exception of the European Union and Japan, all regions are subject to an incentive to leave a coalition that has agreed upon the globally optimal GHG abatement policy. Of course, there are very many uncertainties regarding the physical and economic aspects of GHG emission and global warming. Therefore, the specific numbers of this study quantifying the costs and benefits of GHG abatement in the 12 regions cannot be really taken for granted (as the authors of this study are well aware of and concede). Still, the study illustrates the fundamental difficulties that impede the international coordination of an effective policy to fight global warming. This study (and other empirical treatments) is able to explain why the *Kyoto Protocol* suffers from so many weaknesses and why it is so difficult to agree upon an effective succeeding treaty.

In the preceding analysis, it has been shown that it is extremely difficult for sovereign countries to agree upon a globally optimally designed global warming treaty and to enforce it. Obviously, the globally optimal contract is a very ambitious goal. Thus, the question arises whether the international legislation process has better expectations if it strives for a somewhat more pragmatic goal. A pragmatic program

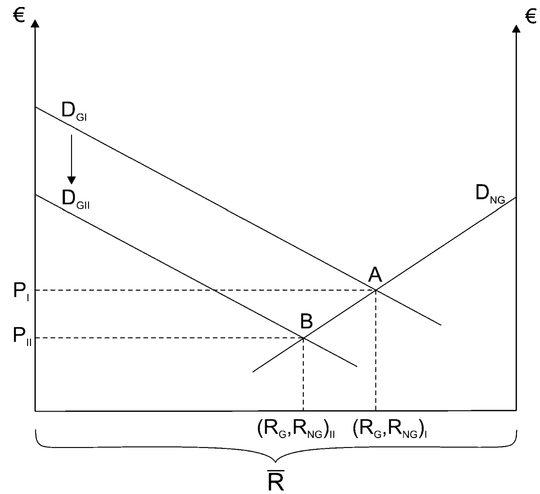
might seek to curb climate change such that the adverse consequences are “tolerable” and allocate the abatement costs among the countries in an “acceptable” way. It has been shown in an extensive literature that this change in the scenario does not lighten up the situation dramatically (See Finus/Caparros (2015) and Finus/Rundshagen (2015) for fundamental issues and recent developments.). Incentive compatible international global warming agreements are likely to be either deep (ambitious reduction target) or wide (high participation rate of countries) and very unlikely to be both.

Leakage

Let us assume that a country or a group of countries, against all odds, would decide to take serious national measures to reduce GHG emission and then fully enforce these measures.

The country or the group of country could take a number of steps that would aim at reducing energy consumption and especially at replacing fossil energy sources with regenerative ones. However, these activities would result in a lower demand of fossil energy sources and thus, ceteris paribus, in a downward pressure on the prices. The decreasing world market prices would stimulate the demand of fossil energy sources in other countries. There, the increasing usage of these energy sources would, ceteris paribus, lead to a rise of the relevant emission though. This interdependence is called “leakage effect” in economics. It results in the dissipation of the self-limiting activities of pioneer countries. If this effect is anticipated, even the countries that would agree to limit themselves would be less willing to do so than without the leakage effect.

This reasoning can be illustrated by Fig. 2. The abscissa shows the stock of a fossil energy resource R fixed at a level of $\bar{R}$. There are two groups of countries in the stylized cosmos of this illustration, green countries (G) and nongreen countries (NG). The green countries are the ones mentioned in the introductory paragraph of this section. They take effective GHG curbing measures. In the pre-policy situation (the time before the green countries decide and act upon their measures), the demand curves for the fossil energy resource are D_{GI} for the green countries and D_{NG} for the nongreen ones. The market equilibrium price (P_I) and the market equilibrium distribution of the resource stock between the two countries $(R_G, R_{NG})_I$ are defined by the intersection of these two demand curves. The effect of the GHG-fighting policy in the green countries is that their demand for the fossil resource is reduced. In the figure, this is illustrated by the demand curve of the green countries shifting down from D_{GI} to D_{GII}. Accordingly, the market equilibrium shifts from A to B in the figure. In detail, that means that the equilibrium price goes down from P_I to P_{II} and the equilibrium distribution of the resource between the two groups of countries changes from $(R_G, R_{NG})_I$ to $(R_G, R_{NG})_{II}$. In terms of global warming, the result of this analysis is that total consumption of the fossil resource is unaffected by the efforts of the green countries. Thereby, GHG emissions jointly produced with the consumption of this resource are unchanged. The GHG emission “savings” generated by the policy of the green countries are completely used up by the additional consumption of the nongreen countries.



Global Warming, Fig. 2 Leakage in the Resource Market

Of course, this is a very stylized reasoning which might not completely cover what is happening in the real world. But, it still illustrates the

"curse of partial solutions" to the global warming problem.

Moreover, the case in which the leakage effect is "transported" via the market for the resource is only one of several forms this detrimental effect might take on. Other forms are discussed in the literature, e.g., in Rauscher/Lünenbürger (2004).

The Dilemma of Intergenerational Allocation

In the preceding analysis, incentive problems of effective global warming policy have been discussed in a framework focusing on the present generation. However, there is also an intergenerational dilemma of global warming policy. GHG reductions conducted in the present are beneficial not only presently but also for future generations. Thereby, an intertemporal externality is generated. The allocative consequences of this externality are analogous to what has been said in the intragenerational context above: Equilibrium GHG reductions of the present generation are too low compared to globally optimal GHG reductions. In the present context, the "global optimality" must be interpreted to comprise the costs and benefits of GHG reduction in all generations. It is plausible without further elaboration that this intertemporal misallocation occurs if the present generation is oblivious with respect to the welfare of future generations. It may come as a surprise, however, that the intergenerational allocation problem does not vanish even if we assume that the preferences of the present generation are not so self-centered. Let us assume the opposite and take the present generation to be *future-altruistic* in the sense that it takes the welfare of future generations into account during its decision-making processes just as it takes its own welfare into account.

To see the point, let us look again at a single country. If a single country trying to optimize the situation for its own citizens and their descendants decides to use less fossil energy, it preserves the energy resources of mankind and causes less CO_2 emissions. However, it cannot be sure that the aforementioned "target group" will benefit from this saving. It must rather anticipate that the saved resources will be consumed by its contemporaries in the other countries. As far as this does not happen, it is by no means granted that the descendants of the country under consideration will benefit from the savings. Instead, the future generations of another country might consume the resources saved by the country under consideration. By this, the aforementioned leakage effect gets a time dimension. Thus, the incentive to save resources is limited for a country that acts "selectively altruistic" insofar that it only cares for its own children and children's children. This incentive distortion can be compared to the one, where individual self-restraint collapses due to the exploitation of *open access resources*. So the individual fisherman does not heed the advice he should bear in mind that he wants to go fishing again tomorrow: The decision-maker knows that present consumers (and, if anything is left, future consumers) will profit from his self-restraint and his own provision will suffer. This results in a phenomenon well known from the economics of natural resources, namely, the individually rational, but collectively unreasonable depletion of freely accessible fishing resources.

Now let us assume (neglecting any concerns based on economic theory) that not a single country would act future-altruistically but the worldwide population as a whole would seriously think about it. Unfortunately, even such an assumption cannot solve the intertemporal incentive problem sufficiently. The trouble is that a present generation, even if it comprises the whole earth, can never be sure whether the succession of coming generations might benefit from a better world climate because of their self-restraint regarding the consumption of natural resources. It is already the very next generation that is tempted to exploit the resources the present generation has left them and to waste them regardless of the CO_2 emissions. Since the generations are naturally limited to a certain period of time, there is nothing generation 1 can do to protect their future-altruistic sacrifices from the exploitation of generation 2. This *constitutional nonprotection* of investments intended to support coming generations is a substantial obstacle for making the investments

G

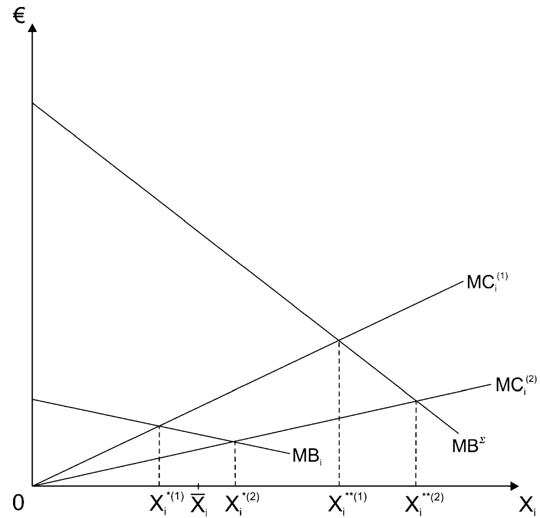
at all. The incentive structure for the present generation is discouraging and can be compared to an investor who thinks about getting economically engaged in a politically unstable country. Since he cannot be sure of how long his investments will be profitable, he is subject to an incentive not to invest at all or only in projects that might be profitable in the short run. Climate protection is a *very* long-term project, though.

The Dismal Science and Its Strive for Climate Change Optimism

Economics is often called the *dismal science*. The pessimistic assessment of the expectations to curb global warming by international GHG policy, as given above, is an obvious tribute to this trademark. However, there are also some strands of economic thinking which might be able to lighten up the picture. Two important examples are *technical change* and *adaptation*.

In the preceding analysis, it has been tacitly assumed that the costs to reduce GHG emissions are given in the sense that they do not change over time. More precisely (and technically speaking), the cost functions mapping alternative quantities of GHG abatement into alternative amounts of cost are invariant. Therefore, in Fig. 1, above, there is only one marginal cost function. This representation is no longer appropriate if we allow for technical change. An important form of this change may be stylized in that with a superior technology more greenhouse gases can be reduced at a given cost than using old-fashioned technology. An equivalent understanding is that a given quantity of greenhouse gases can be reduced at lower cost using the new technology compared to using the old one. In the language of Fig. 1, this kind of technical progress makes the marginal cost curve go down. We illustrate this change in Fig. 3 denoting the marginal cost function, which is valid if the old technology is used, $MC_i^{(1)}$. Accordingly, the cost function illustrating economic circumstances if the new technology is used is called $MC_i^{(2)}$.

It is plausible that technological change making GHG reduction cheaper tilts the scale of



Global Warming, Fig. 3 GHG Abatement and Technical Change

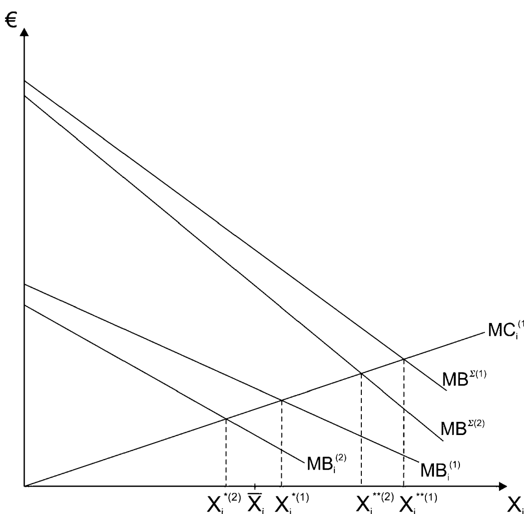
national decision-making into the direction of reducing more greenhouse gases with the new technology than with the old one. In Fig. 3, this is illustrated in that the equilibrium reduction quantity for country i increases from $X_i^{*(1)}$ to $X_i^{*(2)}$. From the perspective of economic theory, this does not reduce the size of the misallocation, since not only the equilibrium quantity increases due to technical change but also the globally optimal quantity. The latter quantity increases from $X_i^{**(1)}$ to $X_i^{**(2)}$. From an ecological perspective, however, technical change is still improving the situation because the higher GHG reduction quantity contributes to mitigate the global warming problem. To illustrate, imagine a GHG abatement target $\bar{X}_i$ which is extra-economically motivated. Then it is possible (and implied in how the level of $\bar{X}_i$ is entered in Fig. 3) that the equilibrium abatement quantity of country i which is valid for the old technology falls short of this extra-economically determined goal, but the equilibrium quantity which holds for the new technology exceeds this goal. Technically speaking, $X_i^{*(1)} < \bar{X}_i < X_i^{*(2)}$ applies.

Another important issue modifying the fundamental analysis presented in sections I–V above is greenhouse gas adaptation. The idea is that the damage associated with a given level of GHG

emissions (a given level of average world temperature increase) is not given, as has been tacitly implied in the preceding analysis. To the contrary, human decision-makers might attenuate the detrimental effects of climate change by prudently adapting to this change. One of many examples is agriculture. The detrimental effect of an increase in the average temperature is reduced (and might even turn to be negative) if the portfolio of crops grown in a certain region is "reoptimized" in the sense that it is adjusted to the changed climate circumstances. Obviously, this is a completely different approach to global warming than that which has been followed in the earlier sections of this essay. Instead of trying to put a break onto the extent (and speed) of climate change, the issue of adaptation is to try to cope with given climate change as well as possible. Of course, the two approaches may be combined.

If, by successful adaptation, the damage of increasing global temperature is reduced so is the benefit of GHG abatement. In terms of the graphical illustrations that have been used above, the benefit curves shift downwards. In Fig. 4, the marginal benefit curves indexed "(1)" denote the situation without adaptation, and the curves indexed "(2)" show the benefits which occur if adaptation is applied. As the curves are drawn in Fig. 4, equilibrium GHG abatement for country i goes down from $X_i^{*(1)}$ to $X_i^{*(2)}$ in the process of adaptation. Globally optimal GHG reduction goes down from $X_i^{**(1)}$ to $X_i^{**(2)}$.

So it is plausible (but very well worth noting because often overlooked in the public discussion) that successful adaptation to climate change attenuates the incentive for GHG mitigation. This is not to be regretted if global welfare maximization is the relevant policy goal. Since adaptation makes global warming less harmful, it is to be welcomed that adaptation tilts the scale of GHG policy decision-making into the direction of less GHG abatement. However, there is a dangerous element in that. This is revealed if we turn from global welfare maximization as a policy goal to the aforementioned goal of achieving an extra-economically motivated standard of GHG abatement. To illustrate this, let $\bar{X}_i$ represent this kind of a standard and assume that $\bar{X}_i$ is "located" between the equilibrium reduction quantity of country i without adaptation, $X_i^{*(1)}$, and the equilibrium quantity with adaptation, $X_i^{*(2)}$. Given these relationships, $X_i^{*(1)} > \bar{X}_i > X_i^{*(2)}$, then adaptation induces policy changes which lead to the result that the extra-economically determined standard is missed. Ironically speaking, one might add: it turns out to be very difficult to take the "dismal" out of the dismal science.



Global Warming, Fig. 4 GHG Mitigation and Adaptation

References

Endres A (2011) Environmental economics – theory and policy. Cambridge University Press, New York

Finus M, Caparrós A (eds) (2015) Game theory and international environmental cooperation. Edward Elgar, Cheltenham

Finus M, Rundhagen B (eds) (2015) Game theory and environmental and resource economics (Part I). Special Issue in Honour of Alfred Endres: Environmental and Resource Economics 62(4), 657–978

Finus M, van Ierland E, Dellink R (2006) Stability of climate coalitions in a cartel formation game. Econ Gov 7:271–291

Rauscher M, Lünenbürger B (2004) Leakage. In: Böhringer C, Löschel A (eds) Climate change policy and global trade. Physica-Verlag, Heidelberg/New York, pp 205–230

Stern N (2007) The economics of climate change: the Stern review. Cambridge University Press, Cambridge

Stern N (2015) Why are we waiting? The MIT Press, Cambridge, Mass./London

Further Reading

Barrett S (2003) Environment and statecraft: the strategy of environmental treaty-making. Oxford University Press, Oxford and New York
Cherry TL, Hovi J, McEvoy DM (eds) (2014) Toward a new climate agreement. Conflict, resolution and governance. Routledge, Abingdon/New York
Finus M, Kotsogiannis C, McCorriston S (eds) (2013a) The International dimension of climate change policy. Special Issue: Environ Resour Econ 56(2), 151–305
Finus M, KotsogiannisC, McCorriston S (eds) (2013b) International coordination on climate policies. Special Issue: J Environ Econ Manag 66(2), 159–382
Nyborg K (2012) The ethics and politics of environmental cost benefit analysis. Routledge, Abingdon/New York

Globalization

Danny Cassimon[1], Peter-Jan Engelen[2,3] and Hanne Van Cappellen[4]
[1]Institute of Development Policy and Management (IOB), University of Antwerp, Antwerp, Belgium
[2]Utrecht School of Economics (USE), Utrecht University, Utrecht, The Netherlands
[3]Faculty of Business and Economics, University of Antwerp, Antwerpen, Belgium
[4]Institute of Development Policy (IOB), University of Antwerp, Antwerp, Belgium

Definition

Globalization is the ongoing process of increasing interconnectedness between cross-boundary actors, driven by flows of people, ideas, goods, and capital. Globalization reduces the relevance of these national boundaries and stimulates the emergence of complex networks that foster the exchange and integration of technologies, economies, governance, communities, and culture.

Introduction

In this contribution, we first define and conceptualize globalization by distinguishing between de jure and de facto forms of globalization and by highlighting its main dimensions. From this conceptualization, we then derive and discuss how we can and are measuring globalization, again in its crucial dimensions; using these measures, we analyze and discuss how globalization has evolved over time and across countries or country groupings. We then look at the consequences of globalization and ways to govern it, from the perspective of three complementary perspectives, the market-centered, state-centered, and people-centered perspectives.

Conceptualizing Globalization

De Jure and De Facto Globalization

An important additional distinction to make both conceptually and empirically is the distinction between de jure and de facto globalization (see e.g., Kose et al. 2009). De jure globalization is expressed in the legal sphere and can be measured through the intensity of regulations, restrictions, and controls on, for example, international trade, capital flows, communications, or the ratification of international treaties. De facto globalization on the other hand measures the actual flows of knowledge, ideas, goods, and people. Both concepts are not necessarily high correlates of each other as high de jure globalization does not automatically lead to high de facto globalization. In certain instances, deliberate repression of de jure globalization may still lead to observed de facto globalization, as ways to circumvent legal restrictions can be found. In the same way does an environment with relatively low de jure restrictions not necessarily stimulate high actual flows. It is thus important when researching or analyzing globalization to be explicit about the kind of globalization one is referring to, while acknowledging that combining both views offer a more complete perspective.

Dimensions of Globalization

Next to distinguishing between “written” de jure globalization and “actual” de facto globalization, it is insightful to disaggregate the different dimensions of globalization. One of the most obvious

transboundary connections is trough international trade. Together with financial connections such as FDI, these constitute the two main categories of economic globalization.

The second dimension of globalization can be found in the political sphere. Political globalization describes the process of an emerging transnational political system that can be observed by the involvement of nations in international organizations, the adherence to international policy-making or political presence beyond the national borders, such as embassies or international NGOs.

The third dimension of globalization is social globalization. Social globalization describes the interconnectedness of people across national or regional boundaries. This interconnectedness can express itself through the ease of physical movement, both temporary and permanent, but also through the interconnectedness via social and communication media.

The fourth dimension, often incorporated in social globalization, is cultural globalization, which describes the process of increasing transmission and assimilation of values, lifestyle, culture, and consumption across boundaries.

Historically, globalization and the interconnectedness between nations was often characterized by asymmetrical North-South relations, especially in terms of trade and labour. However, South-South globalization is increasingly gaining importance, mainly due to the fast growth of China and India and an increase in regional trade agreements in the Global South (Mansoob Murshed et al. 2011).

Measuring Globalization

The KOF Index of Globalization

The KOF Index is the most often used index of globalization. A detailed review of over 100 empirical studies using the KOF index is given by Potrafke (2015), and a list of studies using the KOF index can be found at http://globalisation.kof.ethz.ch/papers. Starting from the premise that globalization is a broad and multi-faceted concept, the KOF index incorporates a wide range of variables in its index, divided into an economic, social, and political dimension (see Table 1). The KOF index was released in 2006 and consequently updated from 2007 on. The latest update, released in January 2018, introduced a considerable amount of innovation. The most important change is the introduction of the distinction between de jure and de facto globalization, an important distinction that is increasingly studied in the literature (see for example Quinn et al. 2011). The previous versions already used variables that represented a mix of both de facto and de jure measures, but no explicit distinction was made.

The KOF index is available yearly for the period 1970–2015. The index as well as each subindex is a score ranging from 0 to 100 with 100 indicating the maximum level of globaliztion, following the procedure of panel normalization. The weights for each variable are calculated using principal component analysis over a period of 10 years, while the weights on the subindices remain the same for the entire panel. Missing values are calculated using linear interpolation or substitution with the closest value. For more details and information on the definition, the dimensions and the calculation of the KOF-index, see Gygli et al. (2018).

De Facto Versus De Jure Globalization

An analysis of the de jure side of the index versus its de facto counterpart shows that it is indeed an insightful distinction to make. When looking at overall globalization for the entire sample, we see that after moving closely together for the first 20 years of the time series, from the nineties on de jure globalization started to increase at a higher rate than de facto globalization (Fig. 1, panel a). When looking at the three dimensions separately, we see diverging trends in the de facto and de jure counterparts (Fig. 1, panel b–d). A common trend in all panels of Fig. 1 is the considerable flattening of globalization in the last decade, possibly due to the lingering effect of the financial crisis. For some subindices, there is even a slight downward trend to be identified, which is typically coined as “de-globalization.”

To see that de facto and de jure concepts are not necessarily high correlates, one can look at the trend of financial globalization (a sub-index of economic globalization, see Table 1) over the

Globalization, Table 1 2018 KOF Globalization Index: Structure, variables and weights. (From The KOF Globalization Index – Revisited, KOF Working Paper, No. 439 by Gygli et al. 2018)

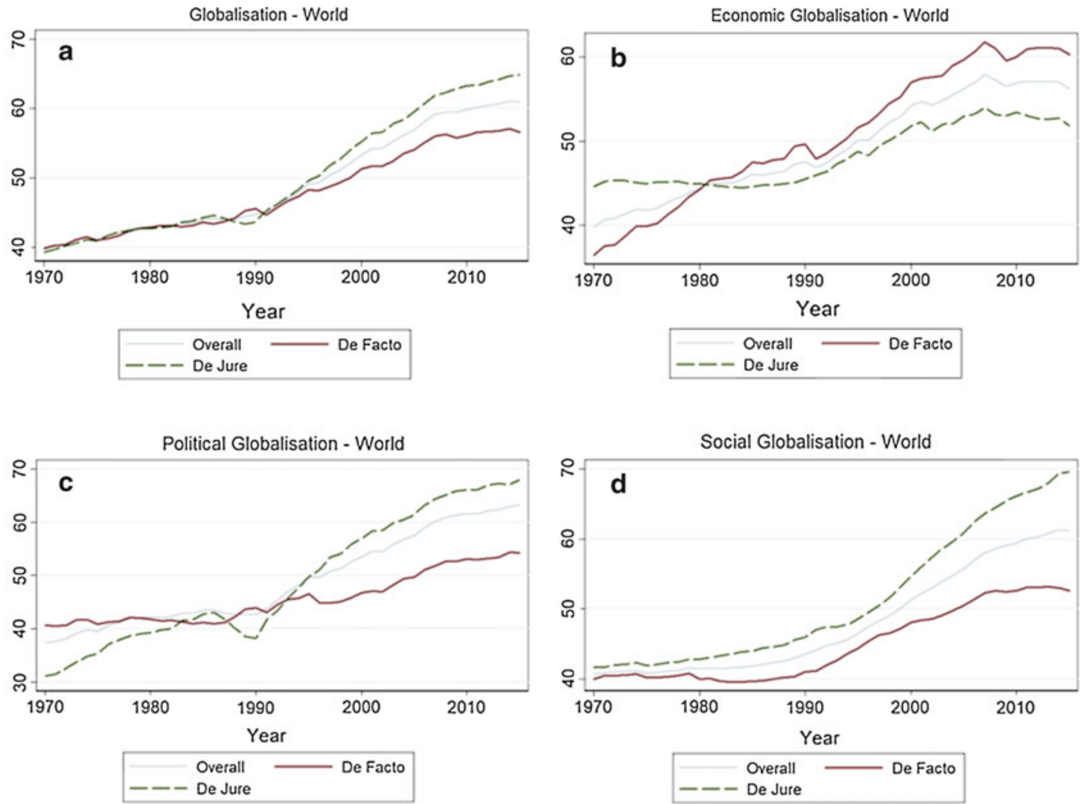
Globalization Index, de facto	Weights	Globalization Index, de jure	Weights
***Economic globaliztion,* de facto**	***33.3***	***Economic globalization, de jure***	***33.3***
Trade globalization, de facto	*50*	*Trade globalzation, de jure*	*50*
Trade in goods	40.9	Trade regulations	32.5
Trade in services	45.0	Trade taxes	34.5
Trade partner diversification	14.1	Tariffs	33.0
Financial globalization, de facto	*50*	*Financial globalization, de jure*	*50*
Foreign direct investment	27.5	Investment restrictions	21.7
Portfolio investment	13.3	Capital account openness 1	39.1
International debt	27.2	Capital account openness 2	39.2
International reserves	2.4		
International income payments	29.6		
Social globalization, de facto	***33.3***	***Social globalization, de jure***	***33.3***
Interpersonal globalization, de facto	*33.3*	*Interpersonal globalization, de jure*	*33.3*
International voice traffic	22.9	Telephone subscriptions	38.2
Transfers	27.6	Freedom to visit	31.2
International tourism	28.1	International airports	30.6
Migration	21.4		
Informational globalization, de facto	*33.3*	*Informational globalization, de jure*	*33.3*
Patent applications	35.1	Television	25.2
International students	31.2	Internet user	31.9
High technology exports	33.7	Press freedom	13.2
Internet bandwidth	29.7		
Cultural globalization, de facto	*33.3*	*Cultural globalization, de jure*	*33.3*
Trade in cultural goods	22.6	Gender parity	31.1
Trademark applications	13.3	Expenditure on education	30.9
Trade in personal services	25.6	Civil freedom	38.0
McDonald's restaurant	23.2		
IKEA stores	15.3		
Political globalization, de facto	***33.3***	***Political globalization, de jure***	***33.3***
Embassies	35.7	International organisations	37.0
UN peace keeping missions	27.3	International treaties	33.0
International NGOs	37.0	Number of partners in investment treaties	30.0

Notes: Weights in percent. Weights for the individual variables are time variant. Depicted are the weights for the year 2015 Overall indices for each aggregation level are calculated by the average of the respective de facto and de jure indices

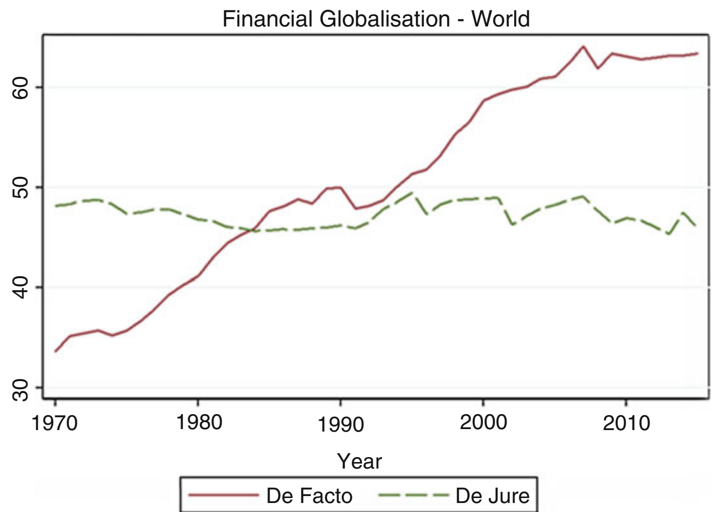
entire period for the entire sample. Figure 2 shows that the de facto financial globalization (measured as FDI, portfolio investment and international debt, reserves and income payments) rose at a high rate over the time period, almost doubling from a score of 34 in 1970 to 63 in 2015. The de jure counterpart (measuring investment restrictions and capital account openness), however, did not exhibit a clear trend, and even decreased slightly over the period from 48 to 46. This example illustrates the importance of the choice of certain variables for measuring and analyzing globalization.

The KOF Index for Different Income Groups

Looking at globalization per income group shows that there is a clear correlation between income level and globalization: the curves retain the same relative position, with the higher income group having a higher globalization score (Fig. 3). This is in contrast with a comparison based on regional

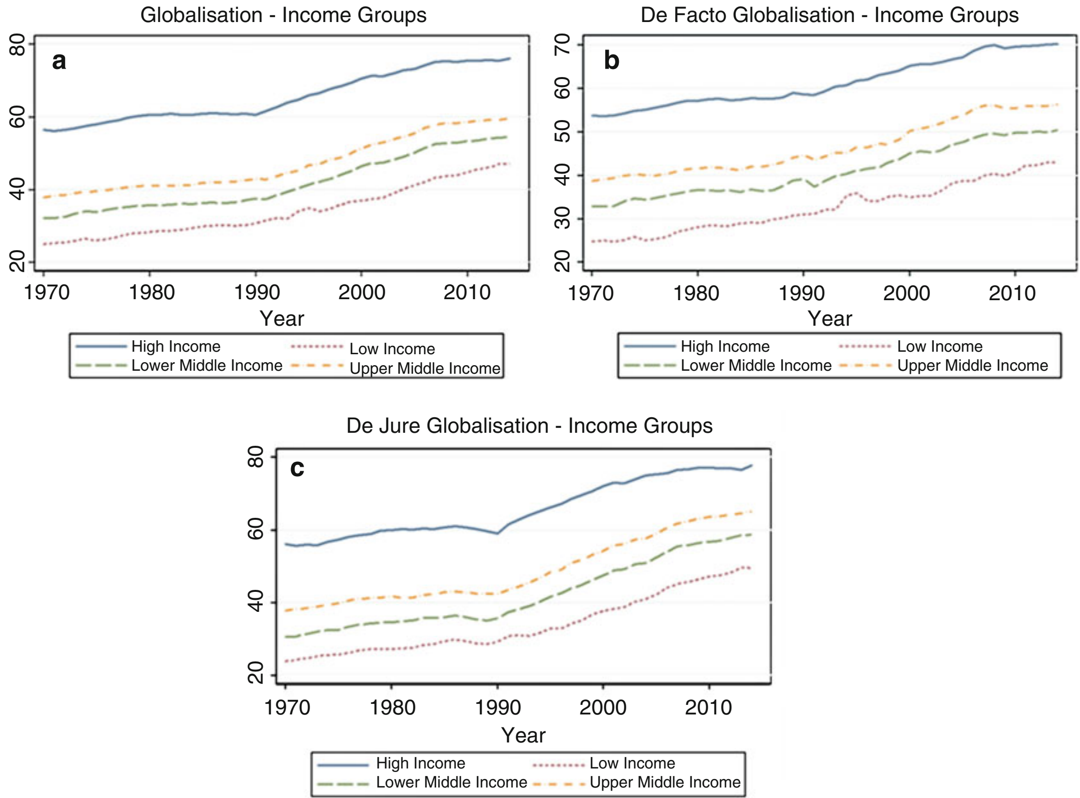


Globalization, Fig. 1 KOF de facto, de jure and overall globalization indices, 1970–2015 – World average

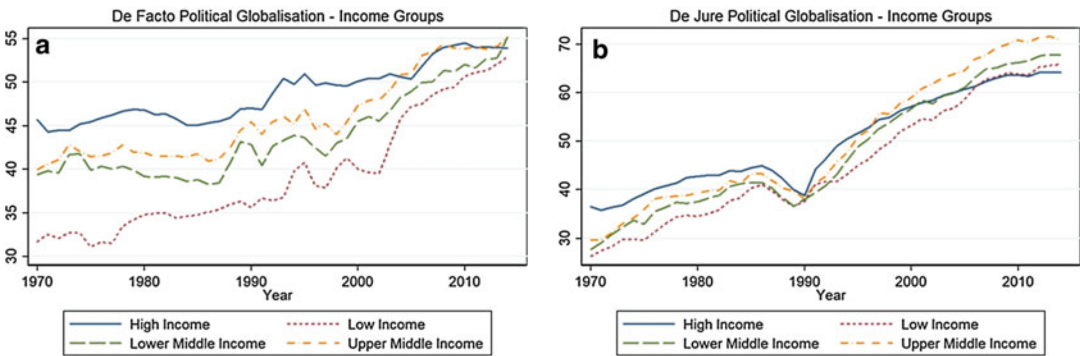


Globalization, Fig. 2 KOF financial de facto and de jure globalization indices, 1970–2015 – World average

groups, for which the curves often look somewhat messier, and exhibit a less consistent ranking. This is possibly due to the influence of particular regional events. However, richer regions are often on the higher end of the globalization spectrum.



Globalization, Fig. 3 KOF overall, de facto and de jure globalization indices, 1970–2015 – Income group averages



Globalization, Fig. 4 KOF de facto and de jure globalization indices, 1970–2015 – Income group averages

In Fig. 3, panel a, we see that overall globalisation per income group rose steadily over the observed timespan, with a slight increase in the growth rate starting from the nineties. When looking at the different dimensions of globalization per income group, we see again that the subdivision between de jure and de facto is useful, as they often paint a somewhat different picture. An interesting example of this is political de facto and de jure globalization (Fig. 4). When looking at de jure political globalization, we see that the different income groups behave

fairly similar, exhibiting a strong growth over the given time period, with the exception of a drop occurring at the end of the eighties but recovering only a few years later. When looking at the de facto political globalization however, we see that although convergence is present, the different income groups start from a far more diverging pattern in the beginning of the time period.

Other Measures

A few other measures have been developed to measure globalization. As such there is the CSGR Globalization Index (Lockwood and Redoano 2005), the GlobalIndex (Raab et al. 2008), the Kearney/Foreign Policy index, and the Maastricht Globalization Index (Figge and Martens 2014). These indices as well as the KOF indices, take the national boundaries as its unit of measurement. This is in contrast with the Person-Based Globalization Index (PBGI) proposed by Caselli (2012). A person-based index counters overemphasis of small countries as well as underemphasis of bigger countries. However, it should be noted that weighting by population is more insightful for some dimensions than for others. Social globalization for example might be interesting to measure on a headcount basis, while for political globalization, which depends largely on national policies, this would make less sense.

Perspectives on Globalization and Its Consequences for Governance

In order to describe the consequences of globalization and ways to try to "govern" or manage it, it is useful to distinguish between at least three different perspectives on globalization, or processes that are triggered by it, described by, e.g., Woods (1998, 2000) as market-centered, state-centered, and people-centered.

The "Market-Centered" Perspective of Globalization

The "market-centered" perspective describes a process of intensification of cross-border exchanges of technology, goods, services, finance, ideas, and ownership, led by multinational enterprises organized in global production and distribution chains, which erodes national markets, making virtually all goods and services becoming essentially "tradables' (as opposite to non-tradables, i.e., goods or services that are only produced and consumed in the local market) and leads to (the "triumph of") global markets. Essentially, it focuses on the perspective of globalization as leading to the global expansion of capitalism. From this perspective, a quite positive prognosis on the results of globalization is derived: the demolition of de jure barriers, the spread of technological advances, combined with efficiency-maximizing behavior of global firms leads not only to long-term production and consumption benefits, and thus higher welfare, but also points at the new opportunities that open up for firms, workers, and consumers all over the world, and particularly also in less developed countries and regions, to gain from being inserted and/or upgraded in these global production chains (Bhagwati 2005; Wolf 2004).

At the same time, while increasing overall welfare, it is clear that this process will produce not only winners but also losers, and that benefits and the costs of globalization are often indeed incurred asymmetrically, necessitating the necessity and ability for complementary measures to be added in an effective way in order to compensate these marginalized and losing ones and try to reintegrate them in a valid and sustainable way in the global capitalist system. In essence, critics in general, and the anti-globalist movement in particular, heavily question the ability of these mitigating effects to be included and, when indeed included, to be effective, and/or, more generally, heavily criticize the ability of global capitalism to provide long-term benefits in a sustainable way (see, e.g., Della Giusta et al. 2006 for a collection of critical writings). A less extreme criticism would suggest positive effects for welfare at modest increases of globalization, that would slowly bottom out at higher levels, and may even become negative at "hyper-globalization" level, hinting that there may something as "too much globalization," as too much of a good thing (for recent

evidence see, e.g., Lang and Mendes Tavares 2018).

A detailed overview of studies analyzing the effect of globalization focusing mainly on this more economic perspective can be found in Dreher et al. (2008) and, more recently, Potrafke (2015). The reviewed studies measure the influence of globalization on macroeconomic performance such as the size of the government, economic performance, labor markets, and capital markets, to name a few. These analyses are mostly performed on a specific geographical unit and as a consequence findings often differ accordingly. When looking at the world as a whole, the general accepted view is that globalization has a net positive effect (Dreher et al. 2008; Wolf 2004). However, the assessment of the effect of globalization on the world as a whole is ultimately a normative question, as the benefits and the costs of globalization are indeed often incurred asymmetrically. As such, studies like Wade (2004) and Dreher and Gaston (2008) conclude that globalization increases income inequality. Goldin and Reinert (2012) research in depth the complex relationship between poverty and globalization and condition the positive effect of globalization on the presence of compensatory policies or mechanisms that mitigate the effects on the actors on the losing end.

Clearly, this market-centered, somewhat "economistic" perspective is too narrow, as if once market forces are unleashed, institutions and policies will follow (Woods 1998), nation-states will see their role diminish and even evaporate and people will become global consumers, being inserted in global labor chains and sharing global values. In other words, state-centered and people-centered perspectives provide not only reactive forces but also complementary insights.

Adding State-Centered and People-Centered Perspectives of Globalization

A state-centered perspective may acknowledge that indeed, the power of nation-states may reduce in some ways as some policies may become less effective in global markets, or some state capacities, such as the ability to tax, may erode, but several counterbalancing forces occur. First of all, as highlighted in the de jure globalization concept, the global integration process has been driven by explicit state policy choices, especially in Western countries. Secondly, states have been using these policies also to protect their citizens, and particularly also the losers of globalization among them, against the negative effects of globalization. Thirdly, in the absence of global markets' ability to make good rules and regulations themselves, it requires states to come together to make and enforce these global rules and regulations. As such, rather that eroding state sovereignty, it mainly transforms it, by shifting some power to a supranational level, be it global or regional.

The need for this cooperative behavior at global level has been enhanced by globalization itself, as more of the pressing challenges of the world, such as environmental degradation, crime, and terrorism or "illegal" migration, have become globalized. As such, actions to solve them have become international public goods, necessitating interventions, such as treaties, new agencies, or regulations at regional or global level (e.g., Kaul et al. 2003). In this new globalized world order, it is important to see who exactly determines the rules in these new supranational forms of cooperation, avoiding a race to the bottom where the weakest link dominates the quality of rules. Moreover, it is also important then to monitor the extent to which less-powerful countries are amply represented or able to voice their viewpoints and concerns at the global table; in other words, what is the quality of global governance (see, e.g., Claes and Knutsen 2011).

The focus on these more nation-state centered perspectives has also shown that the evolution towards global markets did not necessarily lead to a globalized consensus on what constitutes good policies. To give one example, the "Washington consensus" has been rather successfully, leading to the gradual acceptance of more heterodox policies, also in emerging and developing countries, and increasingly being rivalled by the Peking one. And more prominently, globalization has not gone hand-in-hand with the worldwide introduction, at nation-state level, of Western-style forms of democracy.

A "people-centered" perspective adds additional insights, complexities, and counteractions, as it does not only impact people's standards of living but also their values and the opportunities to voice them (see, e.g., Whalley 2008). Again, the proposition that (market-centered) globalization would also lead to the globalization of Western values, and a global culture, has been counteracted by strong upsurges of national and/or religious identity reassertion. It did not lead to a smoother introduction of worldwide human rights values, nor did it realize successful bottom-up promotion of democracy, from the grassroots to nation-state levels. More importantly, a people-centered perspective depicts intensified attention to the asymmetric gains of globalization and the voicing of those left behind in the process. As already stated by Singer and Wildavsky (1993), it is producing a polarized world with "zones of peace" and "zones of turmoil"; these zones of turmoil not only reflect places with a high concentration of losers of globalization in a more narrow sense but also more broadly speaking, zones than may encompass whole regions or countries that are in conflict, or post-conflict, where the state has imploded, and non-state actors have taken over basic service delivery, through all kinds of forms of hybrid governance. Clearly, this will produce increased global inequality. At the same time, increased globalization has also facilitated voicing of concerns and protest against this situation, and also facilitated attempts to escape from these zones of turmoil to the peaceful ones, leading to increased migration moves across the globe, which at the same time start to challenge the very fundamentals of the "market-centered" perspective of globalization. In its most recent form, it is also leading to an increasing wave of forms of populism also in core Western countries such as in Europe and the USA, albeit of different faces, that again threaten to challenge the core fundamentals of market-centered globalization (Rodrik 2017).

This more complex interlinkage between globalization, nation-state power, and democracy is conceptually shown by Rodrik (2011) in his so-called globalization paradox or trilemma, hinting at the fact that we cannot simultaneously pursue democracy, nation-state self-determination, and hyper-market-centered globalization, advocating for a system of "smart globalization" in which we deliberately limit this market-centered globalization, in order to be able to have more political (nation-state) and social (democratic) support from and for those it is supposed to help.

Cross-References

▶ De Jure/De Facto Institutions
▶ Economic Development
▶ Economic Integration
▶ Public Goods

G

References

Bhagwati J (2005) In defense of globalisation. Oxford University Press, Oxford

Caselli M (2012) Trying to measure globalisation, experiences, critical issues and perspectives. Springer Briefs in Political Science. Springer Netherlands, Dordrecht

Claes DH, Knutsen CH (2011) Governing the global economy. Politics, Institutions and Economic Development. Routledge, London and New York

Della Giusta M, Kambhampati U, Wade RH (eds) (2006) Critical perspectives on globalization. The globalization of the world economy, vol 17. Edward Elgar, Cheltenham, 656p

Dreher A, Gaston N (2008) Has globalization increased inequality? Rev Int Econ 16(3):516–536

Dreher A, Gaston N, Martens P (2008) Measuring globalisation: gauging its consequences. Springer Science & Business Media, New York

Figge L, Martens P (2014) Globalisation continues: the Maastricht globalisation index revisited and updated. Globalizations 11(6):875–893

Goldin I, Reinert K (2012) Globalization for development. Meeting new challenges. Oxford University Press, Oxford, 337p

Gygli S, Haelg F, Sturm J-E (2018) The KOF globalisation index – revisited, KOF working paper no. 439

Kaul I, Conceiçao P, Le Goulven K, Mendoza R (eds) (2003) Providing global public goods. Managing globalization. UNDP, New York, 646p

Kose A, Prasad E, Rogoff K, Wei SJ (2009) Financial globalization: a reappraisal. IMF Staff Papers 56(1), pp 8–62

Lang V, Mendes Tavares M (2018) The Distribution of gains from globalization, IMF working paper WP/18/54

Lockwood B, Redoano M (2005) The CSGR globalisation index: an introductory guide. Centre for the Study of Globalisation and Regionalisation working paper 155(04), pp 185–205

Mansoob Murshed S, Goulart P, Serino L (eds) (2011) South-south globalization. Challenges and opportunities for development. Routledge studies in development economics, vol 90. Routledge, London/New York, 348p
Potrafke N (2015) The evidence on globalisation. World Econ 38:509–552
Quinn DP, Schindler M, Toyoda MA (2011) Assessing measures of financial openness and integration. IMF Econ Rev 59(3):488–522
Raab M, Ruland M, Schönberger B, Blossfeld HP, Hofäcker D, Buchholz S, Schmelzer P (2008) GlobalIndex: A sociological approach to globalization measurement. Int Sociol 23(4):596–631
Rodrik D (2011) The globalization paradox: democracy and the future of the world economy. W. W. Norton, New York, 346p
Rodrik D (2017) Populism and the economics of globalization, NBER working paper no. 23559
Singer M, Wildavsky A (1993) The real world order: zones of peace, zones of turmoil. Chatham House, Chatham
Wade RH (2004) Is globalization reducing poverty and inequality? World Dev 32(4):567–589
Whalley J (2008) Globalisation and values. World Econ 31(11):1503–1524
Wolf M (2004) Why globalization works. Yale University Press, New Haven, 398p
Woods N (1998) Editorial introduction. Globalization: Definitions, debates and implications. Oxford Development Studies, 26(1):5–13
Woods N (2000) The political economy of globalization. In: Woods N (ed) The political economy of globalization. MacMillan Press, Houndmills, pp 1–19

Golden Age Piracy

► Piracy, Old Maritime

Good Faith

Hans-Bernd Schäfer[1,2] and Hüseyin Can Aksoy[3]
[1]Bucerius Law School, Hamburg, Germany
[2]Faculty of Law, St. Gallen University, St. Gallen, Switzerland
[3]Faculty of Law, Bilkent University, Bilkent, Ankara, Turkey

We thank Anne van Aaken, Hein Kötz, Alan Miller, Luciano Benedetti Timm, and Reinhard Zimmermann for valuable comments on an earlier draft. All errors are ours.

Abstract

Good faith is a blanket clause under which courts develop standards of fair and honest behavior. It gives ample discretion to the judiciary and entitles a court to narrow down the interpretation of statutes or contracts and even to deviate from codified rules, from the wording in the law or the contract or to fill gaps. The law and economics literature relates bad faith to opportunistic behavior and it is accepted that in specific cases where the application of default or mandatory rules leads to opportunism or where both the law and the contract are silent on risk, the judge can resort to the good faith principle. As a result the parties may delegate part of the contractual drafting to the legal system in addition to having reduced apprehension regarding the possibility of opportunistic behavior from the other side. This allows them to keep contracts relatively short and reduces the costs of defensive strategies.

Definition

Good faith is a blanket clause in civil law under which courts develop standards of fair and honest behavior in the law of obligations, especially in contract law and – to some extent – in the law of property. Courts also define legal consequences resulting from the violations of such standards, such as reliance damages, expectation damages, injunction, imposing a duty of conduct, and invalidation or validation of a contract. The good faith principle entitles a court to narrow down the interpretation of statutes or contracts and even to deviate from codified rules, from the wording in the law or the contract or to fill gaps. Courts use it as a last resort, if and only if the contract itself or the rules of contract law lead to grossly unfair results. "Good faith" also has found its way into jurisdictions other than those of civil law, as well as into the public international law of contracts and international law in general.

Introduction

The roots of the good faith principle can be traced back to Roman times. Even today Cicero's writing

on good faith (De Officiis 3.17 (44 BC)) from more than 2,000 years ago reads as fresh as if it had just been published. It comes close to the modern description of "opportunistic behavior" that economists have developed in recent decades: "These words, good faith, have a very broad meaning. They express all the honest sentiments of a good conscience, without requiring a scrupulousness which would turn selflessness into sacrifice; the law banishes from contracts ruses and clever maneuvers, dishonest dealings, fraudulent calculations, dissimulations and perfidious simulations, and malice, which under the guise of prudence and skill, takes advantage of credulity, simplicity and ignorance" (quoted from Association Henri Capitant and Société de législation comparée 2008). The contemporary content of the principle is not well defined. Its scope is also subject to debate, with some scholars regarding it as a rule of contract law or of the law of obligations and with others relating it to the whole body of civil law.

Objective good faith is an objective standard of behavior, especially for contracting parties. Contract law, with its mandatory and default rules, tries to allocate risk and imposes contractual, pre-contractual, and post-contractual duties and risks fairly and honestly in the way self-interested parties would have chosen had they considered the problem ex ante at the time of contract formation. However, due to the features of a specific case, the application of these legal rules can possess unintended and absurd consequences in individual cases. And even though the black letter rule mostly leads to the desired consequences, it might lead to the opposite in some. In such cases, the good faith principle allows courts to deviate from the black letter rules of contract law or from the contract itself. If the courts find the application and the result of a legal norm in a specific case to be absurd, because they still allow opportunistic exploitation or grossly inefficient risk allocation, they can deviate by using the good faith principle. And they can give the law flexibility if unforeseen contingencies transform fairness into dishonesty and efficient into grossly wasteful allocation of risk. This not only provides adaptability to new circumstances but also saves the parties' transactions costs. More specifically, by knowing that a court has the option to intervene, parties are less inclined to write every contingency into the contract, thus making contracting less costly and defensive measures of parties less necessary.

The principle is not without pitfalls and therefore subject to criticism. Firstly, it gives ample discretion to the judiciary, which can be and has been misused for importing ideology, including totalitarian ideology, into contract law or for promoting private opinions of judges about what they regard as just. Secondly, it might foster judicial activism if the judiciary develops the law proactively in such a way that it replaces parliament. Thirdly, the principle might be inappropriately used to redistribute wealth from the rich to the poor party by adopting a deep pocket approach. Finally, following Hayek, judges who intervene in a contract might not as outside observers have the information for reliably acting in the ex ante interest of honest parties, even if they have the best intention in doing so.

The Good Faith Principle in the Legal Doctrine

The good faith principle is widely used in all civil law countries, for instance, in Germany, Switzerland, Turkey, France, Italy, and the Netherlands. It is also recognized in the US Uniform Commercial Code as well as in unification instruments of private law, including the UN Sales Law, the UNIDROIT Principles for Commercial Contracts, and the Principles of European Contract Law. Good faith is relatively new in American law, but seemingly it also caught on quickly. On the other hand, English law does not recognize a general duty of good faith. In fact, English lawyers think that the courts should only interpret contracts and refrain from allocating risks by way of the good faith principle (Musy 2000; Goode 1992). This does not imply that English contract law cannot correct for absurd consequences of routine decisions, yet the English principles, such as "fair dealing," are narrower than the good faith principle and do not transfer the same amount of discretionary power to the judiciary. Scholars of comparative law have expressed the view that if one looks at how hard

cases of contract law are actually decided in common law and civil law jurisdictions, the differences might be less pronounced than it seems on the surface (Zimmermann and Whittaker 2000). Judge Bingham expressed a similar view with the following words in Interfoto Picture Library Ltd. v. Stiletto Ltd. (1988) 2 W.L.R. 615 (p. 621): (England has no) "overriding principle that in making and carrying out contracts parties should act in good faith " but added that on the other hand "English law has, characteristically, committed itself to no such principle but has developed piecemeal solutions in response to demonstrated problems of unfairness."

A major point of critique of the principle of good faith is its generality and broad scope. This is also in close relationship with the critique that the judiciary can arbitrarily interfere with the contract by using this principle. However, it could be argued that the risk of intolerable legal insecurity and unpredictability is eliminated. In those jurisdictions where the principle is accorded an important role, it is split up into categories and subcategories (Hausheer and Jaun 2003; Grüneberg 2010). Through this subcategorization, an out-of-hand use of the principle by the courts in these jurisdictions is seldom, with each subcategory applying the facts of the case to a series of judge-made tests that possess particular legal consequences on passing the test.

In jurisdictions that recognize a general principle of good faith, there exists a variety of established legal dogmatic forms that define terms and conditions under which the principle of good faith is to be used. Moreover, they also define particular legal consequences if a defendant has violated a standard of behavior. These legal dogmatic forms, all of which are derived from the general good faith principle, provide it with an internal structure, which checks it against willful use. These subcategories are mainly as follows: culpa in contrahendo; contract with protective effects for a third party; liability for breach of trust; adaptation of the contract to changed circumstances (clausula rebus sic stantibus); side obligations from a contract (breach of which constitutes a case of breach of contract); obligation to contract; principle of trust in formation, interpretation, and gap filling of legal transactions; misuse (abuse) of rights; and forfeiture.

The well-established subcategories of good faith substantially eliminate the risk of a judge's arbitrary interference. As a matter of fact, there is a general scholarly agreement on the conditions for resorting to any of the subcategories as well as on the legal results that follow. Still, despite such concrete dogmatic forms, in the jurisdictions where the good faith principle is recognized, it is accepted that such general principles should be used as a last resort, if and only if the formal rules of the contract law lead to absurd consequences. If the good faith principle is a monster, as scholars (Zimmermann and Whittaker 2000) once claimed, it has been domesticated as a farm animal. From an economic perspective, this means two things. First, the good faith principle is a delegation norm for the judiciary. It entitles judges to manually control the law if it does not fulfill its functions. It serves parties if the judges are knowledgeable, loyal, and not corrupt but does a disfavor to them otherwise. Secondly, the development of a tight internal structure is a commitment by the judiciary to credibly signal to the legal community that it applies the principle with self-restraint and self-discipline, ruling out willful, ideological, or biased decisions.

The Good Faith Principle as a Norm to Curb Opportunism

The law and economics literature mainly relates the good faith principle to the prevention of opportunism. According to Summers (1968), good faith is an excluder which "has no general meaning or meanings of its own, but which serves to exclude many heterogeneous forms of bad faith." Burton (1980) relates bad faith to the exercise of discretion by one of the contractual parties concerning aspects of the contract, such as quantity, price, or time. Accordingly, "Bad faith performance occurs precisely when discretion is used to recapture opportunities forgone upon contracting – when the discretion-exercising party refuses to pay the expected cost of performance." Similarly, while defining good faith,

Miller and Perry (2013) refer to the enforcement of the parties' actual agreement. More specifically, the authors argue that the good faith principle protects the reasonable expectations of the parties, which they had while contracting. They criticize however that in the USA in contrast to definitions, courts often resort to community standards, based on what people actually do, and that this can undermine the function of contract. They argue that the good faith principle in US contract law should be based on normative ideas like welfare maximization or the golden rule and not on purely empirical observations. Mackaay (2009) defines bad faith as the legal term for opportunism. Richard Posner describes the principle in a similar way: "The concept of the duty of good faith like the concept of fiduciary duty is a stab at approximating the terms the parties would have negotiated had they foreseen the circumstances that have given rise to their dispute. The parties want to minimize the costs of performance. To the extent that a doctrine of good faith designed to do this by reducing defensive expenditures is a reasonable measure to this end, interpolating it into the contract advances the parties' joint goal" (Justice Posner's opinion from the decision Market Street Associates Limited Partnership v. Frey, 941 F.2d 588, 595).

Elements of Opportunistic Behavior

Williamson (1975) famously defines "opportunism" as "self-interest seeking with guile." In fact, strategic manipulation of information and misrepresentations are to be deemed as opportunistic. The same applies to hidden action aimed at reducing the quality of performance.

According to Dixit, opportunistic behavior is the "whole class of actions that tempt individuals but hurt the group as a whole" (Dixit 2004). Opportunistic behavior is inefficient because it increases transaction costs and reduces the net gain from the contract (Sepe 2010; Mackaay 2011). The risk of opportunism encourages parties to take defensive precautions and write longer contracts to prevent opportunistic behavior as well as the harms that might arise from it. Opportunistic behavior could even incentivize parties to take strict precautions, such as foregoing a contemplated contract altogether. If many contractors were to choose this option, this could destroy entire markets (Mackaay 2011).

Muris (1981) defines opportunism as behavior which is contrary to the other party's understanding of their contract – even if not necessarily against the implied terms of the contract – and which leads to a wealth transfer from the other party to the performer. Cohen (1992) defines it as "any contractual conduct by one party contrary to the other party's reasonable expectations based on the parties' agreement, contractual norms, or conventional morality." Similarly, Posner (2007) defines opportunism as taking advantage of the other party's vulnerability.

Mackaay and Leblanc (2003), who regard good faith as the opposite of opportunism, propose a test to operationalize opportunistic behavior: "an asymmetry between the parties; which one of them seeks to exploit to the detriment of the other in order to draw an undue advantage from it; the exploitation being sufficiently serious that, in the absence of a sanction, the victim and others like him or her are likely substantially to increase measures of self-protection before entering into a contract in the future, thereby reducing the overall level of contracting."

The Necessity to Curb Opportunistic Behavior with a Blanket Clause

The role of contract law can be explained as an endeavor to guarantee the fairness of the contract in the sense of avoiding opportunistic behavior (Posner 2007; Kaplow and Shavell 2002) and through the cost-efficient allocation of risk (Harrison 1995; Schwartz 2003; Schäfer and Ott 2004). In accordance with the purpose of contract law, if the good faith principle were used exclusively to curb opportunistic behavior and to allocate risks in a cost-efficient way in cases in which the established rules of contract law permitted opportunism, it would be regarded as a valuable corrective mechanism giving flexibility and innovativeness to contract law without questioning its rationale for facilitating win-win constellations under fair conditions.

A court can enhance efficiency through the good faith principle in three ways: by (i) not applying a mandatory rule (for instance, by

declaring a non-notarized real estate sales contract as valid), (ii) refraining from the application of a default rule (for instance, by declaring the rule under which partial delivery can be rejected as invalid), and (iii) allocating risks when the law or the contract is silent (for instance, by imposing a non-competition clause) (Schäfer and Aksoy 2015). In the first two cases, the law specifies the risk, but it is clear that the specification of the risk in such a particular case is questionable and gives rise to opportunism. In other words: in such a specific case, applying the norm delivers absurd results, with the blanket clause of the good faith principle on the other hand providing the judge with flexibility that he would otherwise not have had. Finally, both contract and default rules may be silent on a matter consequentially leading to a result that may be neither fair nor cost-saving. In such cases, the principle of good faith can provide efficient risk allocation.

Finally, in line with the legal reasoning, Mackaay (2011) argues that good faith is a last resort tool for preventing opportunism. The author states that there are various anti-opportunism concepts in the law, but when none of these concepts manage to prevent opportunism, the judge will use the good faith principle. Economically speaking, the blanket clause of good faith is a delegation of decision-making authority to the judiciary. If contract law fails to fulfill its aim of curbing opportunistic behavior and allocating risk in a cost-efficient way, thus to guaranteeing fairness and honesty, the judiciary is given the competency for guaranteeing the rationale of contract law, even if this implies deviation from legal rules or from the contract itself. It is obvious that such a delegation norm can work only in jurisdictions in which judges are loyal to the spirit of the law, as well as knowledgeable and non-corrupt.

Good Faith and the “Social Function of Contract”

The self-restriction employed by courts when using the good faith principle has contributed significantly to its international recognition, expansion, and widespread use in many countries and codes and has helped to silence criticism. It is used to maintain the fairness of the contract. It is employed as a last resort; it has developed a rich internal structure, thus preventing willful use. Under these conditions, the good faith principle saves transactions costs, and the business community as well as the whole legal community can regard it as a valuable service.

This proposition comes, however, with a caveat. In some civil law countries, including those of continental Europe, it has been argued that courts have sometimes used the good faith principle to achieve a proper distribution of wealth by interpreting the law or filling gaps in contracts in favor of parties such as consumers, employees, and tenants. Sometimes, such decisions can be interpreted as a correction of market failures, but sometimes they are indicative of a deep pocket approach, meaning redistribution of wealth to the poorer party, even if this neither corrects unfairness nor market failure and is contrary to the terms of the contract itself and to black letter law. This is motivated by distributive justice but comes at the cost of protecting only the insiders not the outsiders and increasing rather than decreasing the costs of using the market making it difficult or even impossible for people without reputation to make promises.

In some countries, the principle is openly used to promote the redistributive goals of social justice, which are usually achieved through the welfare state and statutory laws. In Brazil a case law is emerging in which the good faith principle is not merely used to maintain fairness in the sense of honesty. Under the legal principle of the “social function of contract” the good faith principle sometimes helps to achieve social justice by redistributing income ex post. This can destroy the mutual benefit from a fair contract in which all partners abstain from opportunistic behavior. It seems that such features of “good faith” can be found more often in countries without an effective welfare system or public and private insurance systems. If contracting parties must, however, expect that under certain conditions, in spite of their fairness and honesty, they are stripped from the benefits of the contract ex post by the use of the good faith principle, they might abstain from such contracts altogether. This can lead to dysfunctional markets and considerable costs for

society. The unintended consequences of such interventions, for instance, price ceilings, have been widely discussed in the literature.

Timm (2008) has reviewed some of the Brazilian case law, which spans from not allowing interest on interest, imposing interest rates ex post which are lower than those for government bonds or an injunction requiring an insurance company to cover a surgery or treatment not provided according to the policy, or even preventing an unpaid utility from cutting the supply of electricity. In a lawsuit initiated in Rio de Janeiro, a tenant failed to pay rent for consecutive months. Hence, the landlord sued the tenant bringing a claim for eviction from the property, which the tenant ran as an asylum for the elderly. Despite the Brazilian Landlord-Tenant Law (Law No. 12112/2009) which set forth that a failure to pay rent for consecutive months is a valid reason for eviction, the Appeal Court of Rio de Janeiro ruled for the prolongation of the deadline for an indefinite period of time in order to protect the elderly residents, thus making the landlord a charitable donator. Such decisions help the individual claimant, but not the poor as a group as they deny them the possibility to make credible promises and commitments making market transactions more difficult for them.

According to Nóbrega, the new 1988 Constitution brought about judicial activism by way of establishing a principle- and standard-oriented "social welfare legal order in which the judge had the mission of maintaining social justice (Nóbrega 2012). Consequently, the Brazilian Civil Code regulated both the good faith principle and the social function of contract. Within this scope, the social function of contract, which is derived from the principle of good faith, is regulated by Article 421 of the Brazilian Civil Code as follows: "the freedom to contract shall be exercised by virtue, and within the limits, of the social function of contracts." Although the law refrains from defining the "social function of contract," a prevailing definition of this broad concept is made by Diniz (2007) as a kind of contractual "super-principle," comprising precepts of public order, good customs, objective good faith, contractual equilibrium, solidarity, distributive justice, etc. More specifically, the author argues that the social function of the contract should comprise every constitutionally and/or legally recognized value that might be said to have a "collective" or "nonindividualistic" character.

Conclusion

The law and economics literature relates bad faith to opportunistic behavior, which increases transaction costs, reduces the net gain from the contract, and allows one party to achieve unfair re-distributional gains. Contract law is mainly the endeavor to curb opportunistic behavior and to maintain the cost-efficient allocation of risk. In specific cases where the application of default or mandatory rules leads to opportunism or where both the law and the contract are silent on risk, the judge can resort to the good faith principle. Its use allows for the correction of opportunistic behavior for all those cases in which the black letter law fails to do so. It provides ample discretion to the judiciary and has – from a historical perspective – been misused and heavily criticized. This has not, however, prevented it from becoming an important legal norm in a rising number of jurisdictions and codes.

The good faith principle lost its fearsome features mainly for three reasons. First, its contemporary use is constrained to preserve fairness and honesty among parties, but does not burden honest parties with obligations from community values or social justice in the sense of redistributing income ex post for social purposes. Some exceptions exist however: sometimes camouflaged, sometimes – like in Brazil – easily visible, and more explicit. Secondly, courts use the good faith principle as a last resort when all other routines of judicial decision-making are exhausted yet still lead to absurd legal consequences that allow for opportunistic exploitation or grossly inefficient risk allocation. Finally, the good faith principle has a deep interior structure with subcategories providing routines, tests, and legal consequences resulting from these tests. This makes it difficult for judges to misuse the principle for promoting private or ideological

views about justice and becoming disloyal to the law. Decisions of an individual judge using the good faith principle without observing these self imposed dogmatic constraints are likely to be uplifted by a higher court. These features have led to the increasing confidence that the good faith principle provides a valuable service to parties and enables them to delegate part of the contractual drafting to the legal system. Moreover, it has reduced their apprehension regarding the possibility of opportunistic behavior from the other side. This allows them to keep contracts relatively short and reduces the costs of defensive strategies. In comparison, in jurisdictions which either reject the principle (England) or in which the principle has no long-lasting tradition (USA), there is less delegation and contracts are extensive, are more expensive to draft, and contain long laundry lists of required or impermissible behavior from parties. In the literature, other reasons have been discussed as to why English and US private contract law implies less delegation to the judiciary. This is only one of them.

References

Association Henri Capitant and Société de législation comparée (2008) European contract law – materials for a common frame of reference: terminology, guiding principles, model rules. Sellier European Law Publishers, Munich

Burton SJ (1980) Breach of contract and the common law duty to perform in good faith. Harv Law Rev 94(2):369–404

Cohen GM (1992) The negligence-opportunism tradeoff in contract law. Hofstra Law Rev 20(4):941–1016

Diniz MH (2007) Curso de direito civil brasileiro, vol 3. Teoria das obrigações contratuais e extra-contratuais, 23 edn. Revista e atualizada de acordo com a Reforma do CPC. Saraiva, São Paulo

Dixit AK (2004) Lawlessness and economics – alternative modes of governance. Princeton University Press, Princeton

Goode R (1992) The concept of “good faith” in English law. http://www.cisg.law.pace.edu/cisg/biblio/goode1.html. Accessed 18 Feb 2014

Grüneberg C (2010) § 241. In: Palandt O (ed) Bürgerliches Gesetzbuch. Verlag C.H. Beck, München

Harrison JL (1995) Law and economics. West Publishing, St. Paul

Hausheer H, Jaun M (2003) Die Einleitungsartikel des ZGB – Art. 1-10 ZGB. Stämfli, Bern

Kaplow L, Shavell S (2002) Economic analysis of law. In: Auerbach AJ, Feldstein M (eds) Handbook of public economics, 1661–1784. Elsevier, Amsterdam

Mackaay E (2009) The economics of civil law contract and of good faith. https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/3016/Mackaay_Trebilcock-Symposium%20_3_.pdf?sequence=1. Accessed 18 Feb 2014

Mackaay E (2011) Good faith in civil law systems – a legal-economic analysis. CIRANO – Scientific Publications 2011s-74. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1998924. Accessed 18 Feb 2014

Mackaay E, Leblanc V (2003) The law and economics of good faith in the civil law of contract. https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/125/Article%20papyrus.pdf?sequence=1. Accessed 18 Feb 2014

Miller AD, Perry R (2013) Good faith performance. Iowa Law Rev 98(2):689–745

Muris TJ (1981) Opportunistic behavior and the law of contracts. Minn Law Rev 65(4):521–590

Musy AM (2000) The good faith principle in contract law and the precontractual duty to disclose: comparative analysis of new differences in legal cultures. http://www.icer.it/docs/wp2000/Musy192000.pdf. Accessed 18 Feb 2014

Nóbrega FFB (2012) Custos e Benefícios de um Sistema Jurídico baseado em Standards: uma análise econômica da boa-fé objetiva. Econ Anal Law Rev 3(2):170–188

Posner RA (2007) Economic analysis of law. Wolters Kluwer, Austin

Schäfer HB, Aksoy HC (2015) Alive and well, the good faith principle in Turkish contract law. Eur J Law Econ (in print), https://doi.org/10.1007/s10657-015-9492-1

Schäfer HB, Ott C (2004) The economic analysis of civil law. Edward Elgar, Cheltenham

Schwartz A (2003) The law and economics approach to contract theory. In: Wittman DE (ed) Economic analysis of the law. Blackwell, Maiden

Sepe SM (2010) Good faith and contract interpretation: a law and economics perspective. Arizona Legal Studies, discussion paper no. 10-28. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1086323. Accessed 18 Feb 2014

Summers RS (1968) “Good faith” in general contract law and the sales provisions of the uniform commercial code. Virginia Law Rev 54(2):195–267

Timm LB (2008) Funcao Social Do Direito Contractual No Código Civil Brasileiro Justica Distributiva vs. Eficiéncia Economia. Rev Tribunais, 97, 876:11–43

Williamson OE (1975) Markets and hierarchies: analysis and antitrust implications. Free Press, New York

Zimmermann R, Whittaker S (2000) Good faith in European contract law. Cambridge University Press, Cambridge

Further Reading

Bridge MG (1984) Does Anglo-Canadian contract law need a doctrine of good faith? Can Bus Law J 9(4):385–426

O'Connor JF (1990) Good faith in English law. Darmouth, Aldershot
Piers M (2011) Good faith in English law – could a rule become a principle? Tulane Eur Civ Law Forum 26(1):123–169
Teubner G (1998) Legal irritants: good faith in British law or how unifying law ends up in new divergences. Modern Law Rev 61(1):11–32
Whittaker S (2013) Good faith, implied terms and commercial contracts. Law Q Rev 129(3):451–469

Good Faith and Game Theory

Caspar Rose
Copenhagen Business School, Copenhagen, Denmark

Abstract
This article shows how game theory can be applied to model good faith mathematically using an example of a classic legal dispute related to rei vindicato. The issue is whether an owner has a legal right to his good if a person has bought it in good faith by using updated probabilities. The article illustrates that a rule of where good faith is irrelevant Pareto dominates a rule where good faith may protect an innocent buyer.

Synonyms

Using updated beliefs in modeling good faith and its impact on Pareto efficiency

Definition

If a potential buyer incurs effort and the new updated beliefs result in an increased probability that the good is stolen compared to the initial beliefs, then a buyer is said to be in *bad faith*; if, however, the probability decreases, he is said to be in *good faith*. The buyer is also in bad faith, if he deliberately chooses to stay in ignorance.

The Advantage of Using Game Theory

One may ask, what are the benefits from using game theory for modeling good faith? Game theory relies on a number of assumptions, in particular, the full rationality by the players, which means that the players' objective is to maximize their expected utility; see, e.g., Rasmussen (2004). This means that players make rational choices and have stable preferences.

Game theory has the advantage that it takes into account the players strategic interdependency meaning that the players utility does not only depend on their own actions but also on the opponents actions and vice versa or mathematically expressed: $U_i(a_i, a_j)$ and $U_j(a_j, a_i)$.

Game theory is well suited for analyzing different legal settings, in particular, the incentive effects of different legal rules. Players know what kind of actions is available for them, which naturally is influenced by the legal status, e.g., the possible choice of remedies of breach in a contractual relationship. The players know what the expected outcomes from a sequence of actions are. This is presupposed to be common knowledge among the players. This does not mean that players have perfect information. However, by applying the proper equilibrium concept, one may be able to predict the player's actions, i.e., the outcome of the game. The basic solution concept is the Nash equilibrium where each player plays a best response to each other's actions. This must be refined depending on whether the game is static/dynamic with perfect or imperfect information. In the following analysis, the applied solution concept is the so-called subgame perfect Nash equilibrium which is used for dynamic games with perfect information, but it can include external uncertainty (moves by Nature). In short, game theory is a powerful tool to study games with strategic interdependency where the legal setup defines the player's action spaces; see, e.g., Landes and Posner (1996) for a study of legal rules and strategic interdependency as well as Baird et al. (1998) for an excellent example of how legal rules can be studied using game theory.

The Usefulness of Using Game Theory in Law and Economics

Good faith has been analyzed using game theory, but it turns out that game theory is very well suited for analyzing the economic effects of legal rules. The basic setup of using game theory in law and economics is to study different games under different legal regimes with the same players. To illustrate, it is well known that car drivers and pedestrians behave differently under a regime of strict liability or negligence. Specifically, a car driver may show less precaution under a negligence rule, whereas the pedestrians may be quite careful walking on the sidewalk. In short, law and economics that rely on game theory may be formulated in the following single equation.

$$\Delta \text{Rules/court rulings} \Rightarrow \Delta \text{actions} \Rightarrow \Delta \text{payoffs} \Rightarrow \Delta \text{incentives} \Rightarrow \Delta \text{efficiency}$$

The equation simply displays causal impact of changing the rules or court rulings on efficiency, where Kaldor-Hicks efficiency may be preferred for practical reasons in most cases. Most game theories assume that players are risk neutral, but this assumption may not hold in all cases. Introducing risk aversion in the good faith analysis may extend the analysis to include optimal risk allocation between the parties, including the relevance of insurance, which may serve as an interesting future research topic.

An Illustration: The Sale of Stolen Goods

A person who is offered a television in the parking lot outside a bar is rarely in doubt that the television is stolen. In other situations, a person may find it difficult to determine the likelihood that the offered good is stolen. This is the case when used goods are offered for sale outside private homes (garage sales) or sold through pawnshops. In fact, most stolen goods are traded in workplaces where many people interact and rarely in sinister places. A consumer buying goods in a shop may well presume that the shop seller has legal title to sell the good. However, in many other cases, this is not the case, e.g., especially in auctions on the Internet. There have been numerous cases where a buyer in “good faith” had been contacted by the original owner of a good or artifact, claiming the item handed over. This is a classical legal dispute that goes back to ancient times.

The term *rei vindicatio* stems from Roman law and describes an owner’s right to claim his good back from a person who unlawfully has the good in his possession. *Natural law* or *jus naturale* was the dominant legal foundation for the majority of European countries that coincide with Roman law. The rule was an absolute doctrine of vindication, which was regarded as an inherent attribute embedded in the notion of property rights at that time.

Different legal systems in modern times prescribe various solutions to this fundamental problem of property rights; see Levmore (1987) as well as Mattei (1996) for civil law countries. Some legal systems place the entire risk on the buyer, whereas other systems place the risk on the initial owner. In the latter case, this is the case if the buyer is in so-called good faith. Obviously, the choice of legal rules influence the parties incentives, especially how many resources the initial owner is willing to incur to protecting his property and assets as well as a potential buyers costs of verifying the ownership of the offered good; see Cooter and Ulen (2000).

However, there is a strong strategic interdependency in this situation. A potential buyer, who knows that an initial owner has spent much effort to protecting his assets, implies that the likelihood that the offered good is stolen is small, whereas the contrary is the case in the opposition situation. The key issue is what kind of legal system is most efficient from an economic perspective. The analysis of this question can be analyzed using game theory that includes strategic interdependency between a potential buyer and an initial owner of a good. In fact it turns out that the choice of legal rules in this situation influences parties differently, which has consequences for efficiency.

Modeling Good Faith in the Stolen Good Case

The following results rely in Rose (2010) where the full model is outlined. The article also contains a comprehensive description of how the issue of good faith is treated in several different jurisdictions, including what constitutes good faith. The following is a brief outline of the dynamic game theoretical model with perfect information. There are two players, an initial owner and a potential buyer, who must decide how much effort to spend on protecting his good (denoted c), whereas the potential buyer must determine how much to spend on verifying the ownership (denoted e). Even though there is "perfect information," so the solution concept is subgame perfect Nash equilibrium; there is still randomness which is modeled by introducing Nature as a probability device, denoted $\pi(c)$ which is decreasing in c.

Let $\Pi(c,e)$ denote the potential buyer's "posterior" probability or "updated beliefs" that the good is stolen, given e, which is specified as follows:

$$\begin{aligned}\Pi(c,e) &= \lambda\pi(c) \text{ if } e = e_H \quad \text{where } \lambda \in [0, 1/\pi] \text{ and} \\ &= \pi(c) \quad \text{if } e = 0\end{aligned}$$

These probabilities reflect the idea that a buyer after incurring effort e may revise his initial beliefs that the good is stolen, so that the posterior beliefs that the good is stolen may either increase or decrease. If the buyer incurs effort e and the new updated beliefs result in an increased probability that the good is stolen compared to the initial beliefs, then a buyer is said to be in *bad faith*, i.e., $\lambda\pi(c) > \pi(c)$; if, however, the probability decreases, i.e., $\pi(c) > \lambda\pi(c)$, then he is said to be in *good faith*. The buyer is also in bad faith, if he deliberately chooses to stay in ignorance, i.e., when $e = 0$.

The basic idea is that if it is common knowledge that there is a 20% probability that the good is stolen, but that the posterior beliefs after trying to verify the ownership shows that the good may be equally stolen, then he is said to be in bad faith. The potential buyer is said to be in good faith if the posterior beliefs have shown a reduced likelihood that the good may be stolen.

Furthermore, it is assumed that if a buyer accepts an offer to buy a good that might be stolen this will stimulate criminal activity, since it becomes easier to sell stolen goods; hence, the market for stolen goods becomes more profitable. This will cause more people to enter the market, and as a consequence, the probability of theft increases in the economy; see, e.g., Benson and Mast (2001) for a study on private security services as well as Ayres and Levitt (1996) who show that such crime reduction may generate externalities.

The probability of theft also depends on a potential buyer's decision, i.e., $d = \{accept, reject\}$ with the following specification:

$$\begin{aligned}\pi(c,d) &= \varphi\pi(c,d) \text{ if an offer is } \textit{accepted}, \\ &\quad\ \text{where } \varphi \in [1, 1/\pi] \text{ and} \\ &= \pi(c) \qquad \text{if an offer is } \textit{rejected}\end{aligned}$$

With *a rule where good faith is legally irrelevant,* the good is given back to the initial owner irrespective of any good faith, and the purchaser's utility declines to 0, but at the same time, the initial owner's utility V is restored. It is assumed that the initial owner can always document that she is the rightful owner. In contrast, the good is only handed over to the initial owner under a *good faith rule* if the buyer was in bad faith and if the case is solved by the police. It is assumed that goods cannot be consumed as well as the original owner's utility declines by the value of the good taken away.

We first start up by analyzing what is the optimal situation which afterward is compared with the two regimes, i.e., with good faith and one without good faith.

It can be shown that a regime that does not take into consideration the good faith of a buyer of a stolen good will induce a buyer to accept any offer. The buyer will avoid spending resources on examining the ownership of the offered good. On the other hand, it can be shown that under a rule of law where good faith plays a decisive role in determining property rights, there is a positive probability that a potential buyer will incur costs

in order to verify ownership. Compared to a rule of law where good faith is irrelevant, an owner will spend more resources on seeking to protect his property from being stolen. However, more importantly, a rule where good faith is irrelevant, Pareto dominates a rule where good faith may protect an innocent buyer, as a potential buyer wastes resources on trying to verify ownership and owners of goods incur higher costs in order to deter burglars from stealing their property. This is the case only under the assumptions especially that parties are assumed to be risk neutral. Most game theoretic models assume risk neutrality; hence, payoff functions are simple to work with as they are not concave, as is the case when players are risk averse.

Good Faith in Other Settings

The abovementioned example focuses on a central aspect where good faith may be vital. However, there are other examples such as within tort law. To illustrate, the business judgment rule protects top management from liability, but the legal doctrine has severe incentive effects that can also be analyzed with game theory. According to this doctrine, CEOs and board members are not liable for a decision that results in huge financial losses, as long as it makes the decision on a well-informed basis. Inferior business skills do not constitute a basis for holding an incompetent management liable. A decision which is based on a business judgment made in good faith, but which results in a substantial financial loss for the company, does not make the management liable to legal proceedings by shareholders or creditors.

One may argue that the business judgment rules increase shareholder returns over a situation where management conduct is subject to a strict negligence standard. The rule allows management to engage in more risky projects without the fear of being held personally liable. It is by no means certain that the courts have sufficient knowledge to determine ex post facto, whether a business decision should have been taken or not. In the absence of the business judgment rule, the volume of litigation would increase and transaction costs would rise, not only due to higher fees for lawyers and court fees but also because of increased managerial opportunity costs. The basic idea is that there is a strategic interdependency between managers and shareholders. Good faith protects managers from legal liability, but it also influences the return of the shareholders and their incentives to sue management, which is analyzed in Rose (2011).

References

Ayres I, Levitt SD (1996) Measuring positive externalities from unobservable victim precaution: an empirical analysis of Lojack. John M. Olin Center for Law, Economics, and Business, Harvard Law School. Discussion paper no 197

Baird D, Gertner RH, Picker RC (1998) Game theory and the law. Harvard University Press, Cambridge

Benson BL, Mast BD (2001) Privately produced general deterrence. J Law Econ 44(2), part. 2, pp 725–746

Cooter R, Ulen T (2000) Law and economics, 3rd edn. Addison-Wesley

Landes WM, Posner R (1996) The economics of arts. In: Ginsburgh VA, Menger PM (eds) Selected essays. Elsevier

Levmore S (1987) Variety and uniformity in the treatment of the good faith purchaser. J Leg Stud XVI:43–65

Mattei U (1996) Property rights in civil law. In: Bouckaert B, De Geest G (eds) Encyclopedia of law and economics. pp 162–165, Edwar Elgar

Rasmussen E (2004) The economics of agency law and contract formation. Am Law Econ Rev 6:369–409

Rose C (2010) The transfer of property rights by theft: an economic analysis. Eur J Law Econ 30:247–266

Rose C (2011) Director's liability and investor protection: a law and economics perspective. Eur J Law Econ 31: 287–305

Gordon Tullock: A Maverick Scholar of Law and Economics

Richard E. Wagner
Department of Economics, George Mason University, Fairfax, VA, USA

Abstract

This essay sketches the central features of Gordon Tullock's (1922–2014) contributions to law and economics. My reference to Tullock as a

"maverick scholar" is to indicate that he stood apart from the mainstream of law and economics scholarship as this is represented by Richard Posner's canonical statement that the common law reflects a relentless pursuit of economic efficiency. In contrast, Tullock denied efficiency claims on behalf of common law. Examination of Tullock's contrary claim illustrates how any analytical claim necessarily rests on and is derived from some preceding set of conceptual presuppositions because such qualities as "efficiency" are not objects of direct apprehension but rather are inferences derived from particular theoretical frameworks.

Biography

It is worthwhile noting that Tullock's only academic degree was a JD from the University of Chicago in 1947. He was also an undergraduate at Chicago, enrolling in 1940, and was eligible to receive his baccalaureate degree but elected not to pay the $5 fee required to obtain it because he was already eligible to enter law school and recognized that the law degree he would soon receive would trump a bachelor's degree. His legal studies were cut short, however, by his being drafted into the Army in 1943, where he served in the Infantry and entered into France a week after the D-Day landing in June 1944. After the war, Tullock returned to Chicago to finish his legal program. Tullock spent a few months as a practicing attorney upon graduation from Chicago and then spent 9 years with the US Department of State, including assignments in China, Hong Kong, and Korea. While with the Department of State, Tullock published three papers in major economics journals, all on Asian topics. He had also written a manuscript that was finally published in 1965 as *The Politics of Bureaucracy*, in which he sought to systematize some of his experiences and observations while working with the Department of State. While Tullock recognized that neither legal nor diplomatic practice was how he wanted to live his life, he nonetheless carried into his subsequent academic scholarship the practice-oriented orientation toward his material that his legal and diplomatic activities entailed.

Gordon Brady and Robert Tollison (1991) describe Tullock as a "creative maverick" with respect to public choice theory, and in this designation, they are surely right. Tullock was equally the creative maverick with respect to law and economics. It should be noted that his work in what would generally be considered law and economics is a relatively small portion of his full body of work. This smallness can be seen directly by examining the ten volumes of Tullock's *Selected Works* that Charles Rowley (2004–2006) collected. *Law and economics* was the title of one of those ten volumes. In addition, about 10 % of the first volume titled *Virginia Political Economy* is classified as pertaining to law and economics. The preponderance of Tullock's *Selected Works* was distributed across such fields of inquiry as public choice, rent seeking, wealth redistribution, bureaucracy, social cost, social conflict, and bioeconomics.

Yet we may doubt that any exercise in counting pages can give an accurate representation of Tullock's mind at work. An observer might distribute Tullock's works across discrete fields to provide a scheme of classification, but for Tullock his scholarly activities reflected a mental unity across these superficially diverse fields of inquiry. There was an animating form to Tullock's scholarly inquiries and with that form brought to bear on diverse particular objects of inquiry. With respect to Isaiah Berlin's (1970) essay on Tolstoy, Tullock was surely more the hedgehog than the fox. In his survey of Tullock's scholarship as of the mid-1980s, Richard Wagner (1987: 34) asserts that "someone writing a survey of Tullock's works would surely think he was surveying the work of the faculty of a small university." With respect to his substantive interests, Tullock's scholarship made significant contact with such university programs as public administration, philosophy, history, biology, sociology, criminology, military science, international relations, and Asiatic studies, in addition to his contributions to law, economics, and political science. Despite this substantive variety within Tullock's body of work, his varied inquiries sprang from and reflected a

common analytical core, which I shall explore briefly with particular regard to his scholarship in law and economics.

A. How do law and economics fit together? Melvin Reder (1999) reminds us that economics is an intensely contested discipline. So, too, is law. So how is the compound term "law and economics" to be construed when scholars have a menu of options available to them for thinking about economics and also about law? There are several ways that an economist can theorize about his or her object of interest. A legal scholar faces the same situation.

The primary concept of economic theory is the abstract noun designated as a "market economy." Economists use this idea in explaining how it is that economic activities within a society are generally coherent and coordinated even though there is no person or office who creates that coherence. To be sure, a good number of economists find fault with market outcomes and seek to explain how political power can be deployed to secure what that economist regards as improvement. Nonetheless, the fundamental mystery of economic theory is to explain how societies exhibit coherence and coordination even though each member of that society operates mostly in self-directed fashion. The central answer economists give to explain this mystery relies upon the presence of an institutional framework grounded in the principles of private property, liberty of contract, and freedom of association. This institutional framework brings the legal order directly into the analytical picture, for the efficacy of the market order depends on the quality of the complementary legal framework. Hence, law and economics are two sides of the proverbial coin.

There are, however, different frameworks for economic analysis, and those different frameworks will in turn commend different orientations toward the relationship between law and economics. What Melvin Reder (1982) describes as Chicago economics rests on the twin claims that people maximize given utility functions and that market clear. These claims generate the Pareto efficiency of competitive equilibrium. An economist who thinks that the Pareto efficient character of competitive equilibrium provides a good analytical window for examining economic activity will require a complementary legal framework should that economist choose to explore legal doctrines and procedures. Richard Posner's (1973) animating claim is that the array of common law doctrines and procedures can be rendered coherent by the principle that they reflect the promotion of economic efficiency within society. Posner's claim about law is thus complementary with Reder's claim about the efficiency of competitive equilibrium, when coupled with the further claim that the competitive model is a good analytical window for viewing economic activity within society.

Equilibrium theory is not the only window through which economic phenomena can be viewed. One problem that arises with this approach to economic inquiry is that change comes as exogenous shocks to equilibrium and not as internally generated features of a dynamic economic process. Someone who thinks in terms of the continual generation of change as a quality of an economic system will likewise and necessarily recognize that conflict and also limited and distributed knowledge is part of that system. This alternative analytical window renders claims on behalf of efficiency dubious, mostly because the dynamic economic process is simultaneously extinguishing and creating profit opportunities, with new commercial plans continually being formed simultaneously with other commercial plans being abandoned. Social processes are naturally turbulent and there is no God's-eye vantage point from which efficiency can be determined. This view of social economic processes, which was Tullock's view, will lead in turn to the examination of legal doctrines and procedures through a different analytical window than through which an efficiency-always theorist would use.

James Buchanan (1987) argues that Tullock was a "natural economist." By this, Buchanan meant that Tullock thought naturally in terms of the universal quality of the logic of choice and economizing action, rendering Tullock a theorist in the style of *homo economicus*. There is no small irony in this description, given that one of Buchanan's widely recognized claims is that

economics should be about exchange and not choice. For Tullock, however, the logic of choice was only a point of analytical departure. It was never a destination, for Tullock was always a social theorist who never reduced society to a representative agent. Tullock was a theorist of interaction, perhaps more so than Buchanan, and not a theorist of choice, as Wagner (2008) explains in his comparison of Tullock's *Social Dilemma* and Buchanan's *Limits of Liberty.*

For Tullock it is necessary to distinguish the form by which a theory of utility maximization is expressed from the substantive content of particular instances of human action, all of which are capable of being rendered intelligible within that formal shell. The postulate of rational action means simply that people seek to succeed and not fail in their chosen actions. This is a metaphysical ordering principle and not some behavioral hypothesis. Albert Schweitzer and Adolf Hitler were both utility-maximizing creatures who pursued disparate plans of action that other people appraise to strikingly different effect, but with that difference reflecting the importation of substance into form. George Stigler and Gary Becker (1977) argued that the core of economic theory should rest on the presumption that preferences are invariant across time and place. In contrast, Tullock's analytical core was more like what Ross Emmett (2006) expressed in speculating on what Frank Knight would have said in response to Stigler and Becker. It's also similar to the contrast Richard Wagner (2010: 11–16) advances between neo-Mengerian and neo-Walrasian research programs.

The form of Tullock's theorizing was grounded in rational choice, but for Tullock form was only partly biologically or genetically determined. Tullock would not reduce economics to ethology because continuing and creative social interaction also resided within his analytical core. This brought continuing conflict and continual novelty into his social theory. In other words, Tullock's theoretical framework was one of spontaneous ordering and creative evolutions, as Todd Zywicki (2008) recognizes in his examination of Tullock's critique of the efficiency claims often made on behalf of common law.

I don't recall ever seeing or hearing Tullock refers to Carl Schmitt's (1932) treatment of the autonomy of the political in society, which expressed the idea that politics could never be reduced to law through constitutional design. Yet in a 1965 essay on *Constitutional Mythology* which is not included in his *Selected Works*, Tullock embraced the impossibility of eliminating politics through constitutional law. For Tullock, law and politics were inseparable, and both domains were populated by economizing creatures that gave conceptual coherence to Tullock's various particular lines of scholarship.

B. Tullock and Posner and their Contrasting Orientations. In his contribution for a Festschrift honoring Tullock, Charles Goetz (1987) observes that Tullock's scholarship in law and economics has been largely ignored, in contrast to the reception his scholarship has received in other fields. Tullock (1996: 1), moreover, concurred in Goetz's judgment when he noted that "I have been cast into the outer darkness . . .by the law and economics movement in the United States." In making his point, Goetz compares Tullock (1971) with Posner (1973). Goetz attributes Tullock's lukewarm reception among lawyers as largely a product of Tullock's informal and conversational style of presentation. Recognizing that Goetz wrote this from his vantage point as a professor of law in a premiere law school, I am in no position to challenge that claim. All the same, I think there is more to the differences in reception accorded to those works and their theorists. That difference partly reflects the ability of different conceptual formulations to enable other scholars to do the work they choose to do. Posner (1973) accomplished more than Tullock (1971, 1980a) in this respect. But why is that so? What lessons can be gleaned from this comparison?

Consider one of the several examples where Posner asserted a claim on behalf of legal efficiency and which Tullock (1980b) disputed. Consider the common law principle that railroads owed a duty of care to pedestrians only at crossings, whereas they always owed a duty of care to straying cattle. About this principle, Posner offered the gloss that it would be less costly for a pedestrian to choose a different path than it would

be for a railroad to watch continually for pedestrians. Posner further claimed that the reverse relationship would hold for straying cattle, for it would be more costly for a farmer to fence cattle than it would be for a railroad to keep watch. This gloss fits the efficiency rationale. But is that gloss warranted? Tullock claims that it wasn't. Perhaps walking along the railroad path might save a pedestrian only 100 m as compared with following the sidewalk. But what if that alternative distance were 1,000 m, or even longer? Tullock points out that Posner's empirical claim about relative costs is accompanied by no evidence about those costs. Even more, the cost of watching for humans and cattle are joint and not separable costs. Given that a railroad will be watching for cattle, the marginal cost of watching for humans is zero.

At this point we return to the unavoidable situation where the theoretical framework we use influences what we see in the first place. "Efficiency" is not something that is directly apprehensible. It is rather a conclusion that is drawn from a particular analytical model and with different models often leading to different conclusions. For instance, economists describe a competitive equilibrium as entailing the condition that price equals marginal cost. Cost functions are boundaries that separate what is possible from what is impossible. An average cost function, for instance, defines a boundary where it is impossible for cost to be lower but it is possible for cost to be higher. In recognition of this boundary quality, one might reasonably wonder why economists regularly presume that economic outcomes occur along the boundary and not at some interior position. The answer has nothing at all to do with observation, for it is impossible to observe what is truly the least cost manner of doing something. One might observe that some people produce in lower-cost fashion than other people without being able to determine the lowest possible cost of production. Rather what economists do is make the plausible claim that an individual proprietor would prefer to avoid waste in his or her operation because the proprietor owns the residual between revenues and expenses and transform this reasonable claim into a theoretical generalization.

Tullock's claims that deny the efficiency claims on behalf of common law are not evidentiary-based claims that can point to a failure to use lower-cost options. Tullock's claims are rather advanced in recognition that no evidence has been presented one way or another. In the railroad and cattle case, evidence would entail separately generated data pertaining to different ways of fencing cattle relative to different costs to a railroad of trying to avoid accidents. It may very well be the case that the cost of making finer determinations about costliness in such cases may exceed what people are willing to spend to make such a determination. A judgment will have to be rendered in the case all the same, simply because the case must be resolved in some fashion so that life can go on. This situation doesn't render the common law economically efficient by default, but rather renders the question impossible of being answered.

A famous problem in statistical decision theory involves a lady who claims that in tasting a cup of tea, she can tell whether the tea or the milk was put first into the cup (Neyman 1950). A judge must decide whether to accept or reject the lady's claim. To do this an experiment is designed and evidence generated, from which a judgment is made. With respect to law and economics, this setting is equivalent to asking whether or not the lady is an efficient oracle for determining how a cup of tea is made. Yet there is no way of making such a determination with perfect accuracy because the "truth" of the matter cannot be determined independently of the evidence and standard of judgment used to reach a determination. Errors of judgment will be unavoidable. The lady's claim might be granted when she truly can't tell the difference between the methods. It might also be rejected even though she can tell the difference. Furthermore, an increased effort to avoid one type of error will increase the frequency with which the other type of error is made.

In this setting, efficiency or accuracy is not a binary variable of yes or no but is a quantity that has the property of more and less. Many procedures for reaching a judgment are possible, and these procedures will typically differ both in their costliness and in their ability to avoid one or the

other type of error. Tullock's approach to legal efficiency fits fully within the framework of decision theory. Aside from accuracy and cost, there is also question of the standard of judgment and with higher standards of judgment typically involving more costly procedures. For instance, accuracy will increase as the lady is given more cups of tea to taste, but cost will also rise with the number of cups tasted. It is here where standards of judgment come into play. One judge might feel comfortable granting the lady's claim by invoking a 60 % confidence level, perhaps as expressed by a preponderance of evidence. Another judge might insist upon a 95 % confidence level, perhaps as expressed by overwhelmingly strong evidence. In these situations there is no God's-eye platform from which "truth" can be pronounced. Rather what exists are various decision procedures with varying ability to provide useful evidence and which differ both in their costliness and in their error-avoidance properties.

Tullock's approach to the cost and accuracy of judicial proceedings fits within this decision theory motif. For instance, where Posner analogizes the common law to a competitive market system, Tullock analogizes it to a socialist bureaucracy. There is something to be said in support of each analogy, while there are also things to be said against each of them. The common law system features competition among attorneys for custom and between attorneys at trials. Tullock argues that this system resembles an arms race with both cheap and expensive points of equilibrium and with the process tending toward the expensive equilibrium. On this basis Tullock favors the civil law procedure where judges and not lawyers are the dominant players. But judges are bureaucrats who are paid through tax revenues and not from clients who are seeking their services. Furthermore, law and politics are deeply entangled, as all legal systems are replete with public ordering where the private law principles of property and contract are hemmed in through various politically articulated requirements. Even arbitration, which entails far more private ordering than either common or civil law, requires the willingness of public officials to enforce the judgments that are reached in those proceedings.

Impact and Legacy

A scholar's legacy is, of course, something that will be determined with the passing of time. At the present time, the Goetz-Tullock appraisal of Tullock's reception within the Anglo-Saxon world remains reasonable. That appraisal reflects in good measure the continued predominance of static equilibrium as the core of economic theory. There are, however, signs that various forms of ecological and evolutionary theorizing are gaining momentum within economics. Should that momentum continue to build in the coming years, it is certainly possible that the process-oriented and substantive questions that Tullock addressed will rise within the attention spaces of law and economics scholars relative to formulations that address questions that pertain to models of static equilibrium.

G

References

Berlin I (1970) The hedgehog and the fox: an essay on Tolstoy's view of history. Simon and Schuster, New York

Brady G, Tollison RD (1991) Gordon Tullock: creative maverick of public choice. Public Choice 71:141–148

Buchanan JM (1987) The qualities of a natural economist. In: Rowley CK (ed) Democracy and public choice: essays in honor of Gordon Tullock. Basil Blackwell, Oxford, pp 9–19

Emmett RB (2006) Die gustibus est disputandum: Frank H. Knight's response to George Stigler and Gary Becker's 'die gustibus non est disputandum'. J Econ Methodol 13:97–111

Goetz CJ (1987) Public choice and the law: the paradox of Tullock. In: Rowley CK (ed) Democracy and public choice: essays in honor of Gordon Tullock. Basil Blackwell, Oxford, pp 171–180

Neyman J (1950) First course in probability and statistics. Henry Holt, New York

Posner RA (1973) Economic analysis of law. Little, Brown, Boston. (8th edn, 2011)

Reder M (1982) Chicago economics: permanence and change. J Econ Lit 20:1–38

Reder MW (1999) Economics: the culture of a controversial science. University of Chicago Press, Chicago

Rowley CK (ed) (2004–2006) The selected works of Gordon Tullock. Liberty Fund, Indianapolis

Schmitt C (1996 [1932]) The concept of the political. University of Chicago Press, Chicago

Stigler GJ, Becker GS (1977) De gustibus non est disputandum. Am Econ Rev 67:76–90

Tullock G (1965) Constitutional mythology. New Individualist Rev 3:13–17
Tullock G (1971) The logic of the law. Basic Books, New York
Tullock G (1980a) Trials on trial: the pure theory of legal procedure. Columbia University Press, New York
Tullock G (1980b) Two kinds of legal efficiency. Hofstra Law Rev 9:659–669
Tullock G (1996) Legal heresy. Econ Inq 34:1–9
Wagner RE (1987) Gordon Tullock as rhetorical economist. In: Rowley CK (ed) Democracy and public choice: essays in honor of Gordon Tullock. Basil Blackwell, Oxford, pp 27–38
Wagner RE (2008) Finding social dilemma: West of Babel, not east of Eden. Public Choice 135:55–66
Wagner RE (2010) Mind, society, and human action: time and knowledge in a theory of social economy. Routledge, London
Zywicki TJ (2008) Spontaneous order and the common law: Gordon Tullock's critique. Public Choice 135:35–53

Governance

Paul Dragos Aligica
Mercatus Center, George Mason University, Arlington, VA, USA

Definition

"Governance" is a concept that has emerged to salience at the interdisciplinary interface of several fields relevant for law and economics. A typical governance situation involves not just formal institutions but informal ones, social norms and rules, law, legislation and economic incentives, regulations and enforcement, arbitration and conflict resolution mechanisms, preference aggregation procedures, values and perceptions, and path dependent elements. In the diverse literatures dealing with these phenomena, several major modes of using the notion of "governance" have emerged: governance as describing a recently emerging phenomenon transcending the traditional regulatory, administrative, and public/private choice interface functions of modern states; governance as a functional domain – structures and functions operating in relationship to commons, natural resources and specific collective goods production, management and consumption processes; governance as a theoretical apparatus, a set of concepts, models, and theories aiming at analyzing regulatory, administrative, and public/private choice interface situations, including those depicted by the abovementioned literatures. The concept also applies to the related phenomena associated with the modern corporation and its (legal, political, and social) operational environment. The notion of "corporate governance" is, for instance, a typical example of how this concept is blurring the lines between the "public" and the "private," between regulation and self-regulation, between markets, hierarchies, and networks.

Governance: A Multifaceted Motion

The concept of "governance" has emerged in the last several decades as an alternative mode of theoretically framing the complex political, economic, legal, and social mechanisms and processes associated to the collective decision making, regulation, and administration of political and economic systems, transcending the classical "markets" versus "states" and "public" versus "private" approach. At the same time, it is a way of describing all political and public administration processes and cases taking place in addition to, beyond, or outside of the state/government-centered structures and methods. That covers both the cases of complex modern large-scale advanced postindustrial democracies systems and the cases of groups, small communities, and societies that do not share the scale and complexity of the modern systems. The concept also applies to the related phenomena associated with the modern corporation and its legal, political, and social operational environment. The notion of "corporate governance," is, for instance, a typical example of how this concept is blurring the lines between the "public" and the "private," between regulation and self-regulation and between markets, hierarchies, and networks.

Thus, "governance" is by its very nature, a concept emerging at an interdisciplinary interface. A typical governance situation involves not just formal institutions but informal ones,

social norms and rules, law, legislation and economic incentives, regulations and enforcement, arbitration and conflict resolution mechanisms, preference aggregation procedures, values and perceptions, and path-dependent elements. A large literature (both positive-analytical and normative-prescriptive) has grown under this label with roots in political economy, law and economics, political philosophy, public administration, management studies, anthropology, sociology, and history. At the core of this exploding literature stands a large variety of "governance" forms and domains and their multiple facets, as framed using diverse theoretical lenses. This includes governance as a political theory concept, as a structure (architecture of institutions), as a process (the dynamics of functions), as a mechanism (procedures and decision rules), as a public choice intuitionalism theory, as a strategy (actions related to control and institutional design shaping preferences and decisions), and as a normative ideal.

Due to its multifaceted nature and its versatility and traveling capacity from one discipline and intellectual tradition to another, the notion of "governance" is also increasingly assuming all the features of a typically "essentially contested" concept, gradually joining in this respect other key terms that form our political vocabulary such as "democracy," "liberalism," "government," "politics." "Essentially contested" concepts do not display a mere ambiguity about usage, about a core meaning, which could be solved using standard analytical methods, but reflect a deeper tension or disagreement about different configurations of theoretical, descriptive, and normative elements that are implicit in the usage of the concept by different schools of thought, disciplines, and intellectual traditions. These systems or configurations of ideas, which embed the concept in case, are sharing some elements –sometimes at all three levels, descriptive, theoretical, or normative – but not all of them, a fact which makes the disentanglement of the meaning of the concept in case always extremely difficult and contestable. When it comes to the concept of "governance" that means that the usual caveats applied to any encyclopedia entry dealing with this special class of concepts, should apply to it as well.

In what follows, this entry will illustrate the potential of the family of perspectives gravitating around the concept, by focusing on three modes of using it. All are at the general and encompassing level, all are backed by solid literatures: (1) governance as describing a recently emerging phenomenon transcending the traditional regulatory, administrative and public/private choice interface functions of modern states; (2) governance as a functional domain – structures and functions operating in relationship to commons, natural resources and specific collective goods production, management and consumption processes; (3) governance as a theoretical apparatus, a set of concepts, models, and theories aiming at analyzing regulatory, administrative, and public/private choice interface situations, including those depicted by the above mentioned literatures.

The New Realities of "Governance"

The starting point of the first is the observation that a "new reality" has been created at both the national and the global levels: Globalization, competition, technological change, financial pressure, and increased demands due to changes in perceptions, preferences, values, and expectations put an increasing pressure of governments. A shift has taken place. The new conditions have generated a pressure to adjust the regimes centered on nation-states and on the hierarchical command and control arrangements based on Weberian bureaucracy and Wilsonian public administration. The typical national governmental structures are overextended, lacking capabilities and unable to administer, intervene and control as directly and as effectively as before. Because the state capacity of control diminishes, it has to adjust its modus operandi. A change of approach has to take place focusing now on indirect, negotiated influence, "outsourcing" or building partnership with the private sector and civil society.

Private actors and hybrid forms of organization become players in traditionally public domains, blurring thus the public-sector/private-sector

distinction. The retreat of government's control has encouraged an expanding role of civil society, a diminished view of the role of elected officials, an emphasis on political entrepreneurship, the pooling of public and private resources, and reliance on market or quasi-market discipline. Decentralization and de-monopolization has encouraged a variety of formal and informal hybrid public-private organizational forms, pluralist, multilevel, and polycentric arrangements. These changes have led to a move from politics, hierarchies, and communities to markets; from provision to regulation; from public authority to private authority; from big government to small government; and from regulation based on command and control to flexible and diverse forms of regulation in which self-regulation is an important element. The overall result has been the growth of "governance without government." Given these developments, "governance" has come to indicate an entire range of public and semi-public mechanisms and institutions for implementing policy, working either as alternatives to "government" or in conjunction with it. The boundaries between the public, private, and voluntary sectors have changed and with that the very role of the stat (Peters and Pierre 1998).

A very telling metaphor recurrently used to portray these phenomena is based on the distinction between "rowing" functions (service provision, taxation, redistribution) and "steering" functions (rulemaking, monitoring and enforcement). Expressed in this terms, the point is that a shift has taken place. Initially, states were bent on command and direct provision, insisting on doing both the "rowing" and "steering." Now states are increasingly focused on "steering," while markets, the voluntary sector, and nonprofit civil society do the "rowing." The literature is not displaying a unitary interpretation; there are alternative modes of understanding the "new reality." For instance, on the one hand, there is the "hollowing out the state" thesis whereby authority is shifting either "up" to markets and transnational institutions or "down" to local governments, business communities, and nongovernmental organizations. On the other hand, there is the thesis that the state has in fact reasserted its authority by shifting its focus at the meta-level towards "regulating the mix of governing structures such as markets and networks and deploying indirect instruments of control," the notion of meta-governance (Levi-Faur 2012, p. 36).

As noted, the developments outlined above have generated a robust literature. One could appreciate the publication of volumes like *The Oxford Handbook of Governance* (Levi-Faur 2012) and *The Oxford Handbook of Public Management* (Ferlie et al. 2007) as one of the most significant indicators of the status and impact of that literature. The first volume charts the "new reality" created at both the national and the global levels as the role of the state and the boundaries between the public, private, and voluntary sectors have changed. It illustrates how these themes are at the core of both the academic and applied agenda, and it points out to some ways of theorizing, describing, and analyzing them. The second volume takes the applied angle and documents the emergence to salience of themes such as proactive administration, administrative responsiveness and responsibility, street-level bureaucracy, citizen participation, public-private partnerships, decentralization, hybridity, contracting, and NGOs.

Governance as a Functional Domain

The second mode of illustrating the "governance" perspective is as the study of institutional arrangements dealing with specific social dilemmas of collective action, in particular circumstances defined in functional, local, and contextual terms. Its most salient application is regarding the commons, natural resources, and specific collective goods production, management, and consumption processes. The contribution of 2009 co-recipient of Nobel Prize in Economics, Elinor Ostrom, is illustrative, her "governing the commons" work being a source of inspiration and considered as paradigmatic in this respect (Ostrom 1990).

From this perspective, the notion of "governance" is a term associated to a set of institutional and procedural solutions to collective action situations generated by the collective production, financing, and consumption of goods and services

that have particular properties that diverge from the private goods ideal (in terms of excludability and jointness of consumption) (Ostrom and Ostrom 2004). A "commons" is a good that is shared by multiple users and has two properties: rivalrous and nonexcludability. The use of a commons by one user changes the state of the commons and has implications for the preferences and behavior of other users, who change their behavior and overuse or deplete the resource, leading, without intention, on the long run, to a decrease of the welfare for the entire group: This is the so-called "tragedy of the commons." It is by definition a collective-action problem: The incentives defining a situation induce a divergence between the individual and collective logic, generating a rush to enjoy/use/consume the resource. Thus, the commons is predicted to deteriorate, affecting the welfare of all the commons users.

E. Ostrom and her team of researchers not only built an extensive database and a theoretical and methodological apparatus to study how people deal with such situations, but also engaged in extensive fieldwork examining particular case-studies of governance systems of fisheries, irrigation systems, pastures, forests, etc., in both developed and developing countries, at various levels and scales. At the core of the agenda was a demonstration that the "tragedy of the commons" could be and has been avoided in many cases, without making use of state-based mechanisms and enforcement. Ingenious institutional arrangements and social norms were created and enforced, leading to viable and resilient governance structures. The "governing the commons" research program has thus shown that "governance" is possible outside the state-market dualism that dominated mainstream policy analysis. It illustrates empirically and analytically not only that community-based governance may be a viable alternative, but also that governance functions may be operationalized through diverse institutional forms. Hence the notion of "governance" has come to be associated to the notion of "institutional diversity." In addition to that, a significant part of the work was focused on the conditions and mechanisms motivating the commons users to behave cooperatively. E. Ostrom identified a set design principles, noting that they characterize robust institutions used by communities managing common-pool resources such as forests or fisheries. With that, the notion of "governance" has also come to be associated to the notion of "institutional design."

Last but not least, the "design principles" demonstrated the possibilities for self-governance of human communities and the ability of people to bottom up design and manage solutions to social dilemmas. The Ostroms' research program exploring the governance of the commons has thus to be seen as part of a larger program having an assumed applied and normative dimension: "What is missing from the policy analyst's tool kit – and from the set of accepted, well developed theories of human organization – is an adequately specified theory of collective action whereby a group of principals can organize themselves voluntarily to retain residuals of their own efforts" (Ostrom 1990, p. 24).

G

Governance as a Theoretical Apparatus

The third mode of using the concept of "governance" is precisely in connection to this theoretical toolkit. "Governance theory" is in this respect the range of conceptual and theoretical instruments available for the study of the phenomena identified and described above. From this perspective, "governance" refers to a description of "governance phenomena" (as social coordination, cooperation, and institutionalization) as seen through the lenses of this theoretical apparatus. It is a dual relationship. Theoretical lenses lead to a specific description and diagnosis of governance problems. While the problems, thus defined, entail for their analysis and solutions, the logic and social mechanisms implied by the toolkit. Most of the components of the toolkit have origins in economics: public choice, industrial organization, institutional theory, applied game theory, etc.

The theory of private and public goods, as already suggested above, offers a very effective illustration in this respect: Because the nature of goods determines the practicality of the various institutional structures meant to produce, finance, distribute, and consume these goods, a typology

of goods is a first step in understanding institutional diversity. Hence, a governance theorist approaches institutional arrangements as driven by a functional structure shaped by the nature of goods or services that those institutions are supposed to satisfy. Thus, a typology of goods (based on the "exclusion" and "jointness of use or consumption" dimensions, sometimes further elaborated in game theoretical forms) helps identify typologies of institutional forms. At their turn, the typologies of institutional forms help chart and analyze governance situations. This is the strategy advanced, for instance, by E. Ostrom and V. Ostrom (2004) or by 1986 Nobel Prize in Economics, James Buchanan (1967).

In a similar way, one may use transaction cost analysis, on the lines pioneered by Ronald Coase (1991 Nobel Prize in Economics) and further elaborated by O. Williamson (2009 Nobel Prize in Economics co-recipient). The logic is similar. The diverse nature of the transactions people have to engage into and especially their associated costs (search, monitoring, enforcement etc.) determine the practicality of the various institutional structures meant to deal with those transactions. (Coase 1992; Williamson 1979) Hence, a typology and analysis of transactions and of the costs associated to them is a first step in understanding institutional structure, function, and diversity. At its turn, the typology of institutions helps chart and analyze governance situations.

The list may go on: externalities, agent-principal, rent seeking, asymmetric information are all part of the catalogue or toolbox of the governance theorist or analyst. It is a modular toolkit which has a double function: On the one hand, it offers a vocabulary and thus the lenses to identify and describe governance situations, and at the same time, it offers analytical instruments to explain the social mechanisms and processes in place or to suggest such mechanisms and structures as governance solutions to the problems thus diagnosed. In this respect, it has strong connections with mechanism design theory (2007 Nobel Prize in Economics for Leonid Hurwicz, Eric Maskin and Roger B. Myerson) – an approach that uses "reverse game theory," which starts with a desired function or objective, and tries to identify and reconstruct the governance and institutional mechanism which may be able to achieve that specific objective or realize that functional solution (Maskin 2008). As such the "governance" approach assumes strong and unequivocal normative and applied level dimensions.

Cross-References

- ▶ Bloomington School
- ▶ Constitutional Political Economy
- ▶ Public Choice: The Virginia School

References

Buchanan JM (1967) Public goods in theory and practice: a note on the Minasian-Samuelson discussion. J Law Econ 10:193–197

Coase RH (1992) The institutional structure of production. Am Econ Rev 82(4):713–719

Ferlie E, Lynn LE, Pollitt C (2007) The Oxford handbook of public management. Oxford University Press, Oxford

Levi-Faur D (ed) (2012) The Oxford handbook of governance. Oxford University Press, New York

Maskin ES (2008) Mechanism design: how to implement social goals. Am Econ Rev 98(3):567–576

Ostrom E (1990) Governing the commons: the evolution of institutions for collective action. Cambridge University Press, Cambridge, UK

Ostrom E, Ostrom V (2004) The quest for meaning in public choice. Am J Econ Sociol 63(1):105–147

Peters BG, Pierre J (1998) Governance without government? Rethinking public administration. J Public Adm Res Theory 8(2):223–243

Williamson OE (1979) Transaction-cost economics: the governance of contractual relations. J Law Econ 22(2):233–261

Government

Walter E. Block
Joseph A. Butt, S.J. College of Business, Loyola University New Orleans, New Orleans, LA, USA

Abstract

The institution of "government" is perhaps more debated among philosophers, political scientists, economists, historians, and other

social scientists than any other. In order to give a comprehensive assessment of this concept, I will offer several perspectives on it.

Definition

Government is a criminal gang with a lot of good public relations, so much so that most people think is has legitimacy

The Theocratic

In this vision, the king is the head of government and is also the leading religious figure. Sometimes, this type of emperor is actually God himself. As owner of the kingdom, literally, he is the unquestioned master of not only every stick, blade of grass, and other property in his realm but also the people in it. They are in effect his slaves; he is their master. His authority is unquestioned, both from the religious and secular perspective. One need not spend too much time criticizing this form of government. How can it be proven that any given individual is actually God or his chosen spokesman? Why is it justified that it is from this particular family that kings, queens, and princes emanate? Typically, this is based on success in war, but any justification for this process must rest on the unproven contention that "might makes right."

The Democratic

Here, at least in pure form, the government is elected by the governed, and all decisions are based upon majority rule. The franchise might be broad or very narrow (e.g., limited to white, male property owners). The inclusive version of this system enjoys wide support. Most countries in the so-called civilized world engage in this political form. According to Winston Churchill, "democracy is the worst form of government except for all the others that have ever been tried." But there are problems here. For one thing, Hitler, no popular figure, came to power through a completely democratic process, not via a coup d'état. (Of course, it cannot be denied that he remained in power by abrogating the democratic system.) One is tempted to conclude from this one fact, "so much for democracy." But this is only the tip of the iceberg, for there is such a thing as "the tyranny of the majority." Just because a majority of the electorate prefers one candidate or policy over another does not guarantee he or it is justifiable. In many southern towns in the United States in the nineteenth century, a majority of the people might well have voted to lynch all black people.

The Republic

This form of polity allows for democracy but is subject to a constitution. That is, the society can vote alright, but it cannot cast a ballot for anything and everything. For example, some things cannot be determined by majority vote, such as should a certain minority of people be eliminated. This seems like an improvement over pure democracy, but it all depends upon how good the constitution is and whether or not it is followed. For instance, the US constitution allowed for slavery; the majority need not have voted to enslave black people: the constitution already allowed for that. Another example was the constitution of the Soviet Union. It protected all sorts of minority rights. But this relatively splendid document was all but ignored.

What have libertarians had to say about government? I pursue this question since this philosophical movement has some interesting and important things to say about this institution. Four different versions of libertarianism can be identified in terms of their vision of proper government. They are as follows:

1. Anarcho-capitalism
2. Minarchism
3. Classical liberals, or Constitutionalists
4. Free market supporter

Libertarianism is predicated upon two basic building blocks: the nonaggression principle

(NAP) and private property rights based upon homesteading. The first maintains it is improper to threaten or initiate violence against an innocent person or his property. The second is a theory as to how just property titles come into being. They start with self-ownership, continue with the mixing of one's labor with land and other natural resources, and conclude with any voluntary means of title transfer, such as buying and selling, gifts, and gambling.

Anarcho-Capitalism

The anarcho-capitalist version of libertarianism cleaves most closely to these principles. It brooks no compromise, none whatsoever, with private property rights and the NAP. That is, it concludes that government is a per se violation of rights. All governments do two things: they tax their subjects and forcefully prevent other entities from competing with them, within their geographical territory. But taxation is theft. People did not agree to be taxed. Any "social contract" to the contrary is a figment of the statists' imagination. Taxes are not akin to club dues. The latter rises from contract, not the former. Thus, for the anarcho-capitalist version of libertarianism, government is not merely akin to a robber gang, but indeed takes on that exact description. The only difference between it and a bunch of outright thugs is that the state has legitimacy in the minds of most people and that gangsters do not. So similar are the two, apart from the far better public relations of the government, that when a criminal engages in a crime such as murder, rape, and theft, he can be said to have undertaken a governmental act (apart from the seeming legitimacy of it).

The scholars most closely associated with this perspective are Murray N. Rothbard, Lysander Spooner, and Hans-Hermann Hoppe.

Minarchism

The libertarian perspective consistent with the principles of this philosophy to the second greatest degree advocates minimal state government and is called minarchism. Here, the government is indeed a legitimate institution, but the sole function of the state is to protect the persons and property rights of its citizens. To that end there are three and only three legitimate governmental institutions: One, armies to protect domestic inhabitants against foreign aggression. (Not to become policeman to the world; not to engage in imperialistic ventures here, there, and everywhere; solely for *defense.*) Think in terms of a gigantic, ferocious, and very powerful Coast Guard. The second justified institution is police, to protect us from local violators of the NAP: to stop murderers, rapists, thieves, etc. (not to arrest victimless criminals for "crimes" involving consenting adult behavior such as pertaining to drugs, pornography, prostitution, gambling). Third, courts to determine guilt or innocence of those accused of real crimes and to address contract disputes. Minarchists strictly limit government to these three institutions.

Key contributors to this perspective are Robert Nozick, Ayn Rand, and Ludwig von Mises.

Classical Liberals, or Constitutionalists

Next in the libertarian pantheon in terms of consistency to principle are the Classical Liberals or Constitutionalists. To the three functions of government, advocates of this version of libertarianism would add a few more: dealing with communicable diseases and storms and floods and providing "public goods" such as roads and utilities like gas, electric, water, sewage removal, etc. Some in this category would include anything mentioned in the US constitution, such as post offices and coinage. Ron Paul is perhaps the most famous person connected with this viewpoint.

Free Market Supporters

Here, there are lots of compromises: health, education, welfare, externalities, antitrust policy, and monetary crankism (e.g., rejection of the gold standard). So much so that there are questions as to whether or not this viewpoint can be considered libertarian at all. On the other hand, people associated with this perspective are staunch advocates of the free enterprise system and adamantly oppose government incursions such as rent control, minimum wages, tariffs, corporate welfare, crony capitalism, and entry restrictions for business. They favor the profit and loss system. Those

associated with this point of view include Milton Friedman, Friedrich Hayek, and Rand Paul.

I now move to a discussion of anarcho-capitalism, the purest form of libertarianism, since its indictment of the government is the most accurate. That is, the state really is an organized means of NAP violation. It is the only institution in all of society which may legally initiate violence against innocent people. As such, it deserves intensive analysis, in any effort to understand what government is all about.

Let us start off with Rothbard (1977) who avers that it is important that any libertarian

> . . . hates the existing American State or the State per se, hates it deep in his belly as a predatory gang of robbers, enslavers, and murderers. . . the State – *any* State – is a predatory gang of criminals. . .. the radical (libertarian) regards the State as our mortal enemy, which must be hacked away at wherever and whenever we can. To the radical libertarian, we must take any and every opportunity to chop away at the State, whether it's to reduce or abolish a tax, a budget appropriation, or a regulatory power. And the radical libertarian is insatiable in this appetite until the State has been abolished. . .

There are those who would quarrel with such an analysis and maintain that the government is really a voluntary organization, like a club. And, just as in the case of the golf, or tennis, or chess club, every member must pay his dues. In this case that would be taxation. But Schumpeter (1942, p. 198) puts paid to this excuse for brutality: "The theory which construes taxes on the analogy of club dues or of the purchase of the services of, say, a doctor only proves how far removed this part of the social science is from scientific habits of mind." The point is there is simply no contract, "social," or otherwise that binds people together. Only a very few people signed the declaration of independence. How can it be binding on an entire populace?

But did not three-quarters of the states ratify the national government in the case of the United States, by a majority in each case? This cannot be denied. But how does that bind the people in the states that did not ratify this document? How does it obligate those who did not vote or voted against the proposal? No rational court would hold a person responsible for a commercial contract to which he did not agree, just because others, even many others, did so.

Suppose a slave master allows his property to vote on overseer Goodie, who will only beat them once per week, and overseer Baddie who will do so three times per day. The slaves opt for the former. Does this mean these poor unfortunates have agreed to be slaves and that they agree to have Goodie whip them? Of course not. Spooner (1966[1870]) is justly famous in libertarian circles for making just this point. He also disposed of the argument that tax compliance demonstrates agreement with government overseers. We also pay the highwayman when he holds a gun on us. This hardly provides evidence that our "transaction" with him is a voluntary one, and unless it is the ethical case for the state has no foundation.

What about the following argument: "If you don't like it here, under this government, you are free to leave. The fact that you stay indicates voluntary agreement with these institutional arrangements." Stuff and nonsense. One difficulty with this is that it argues in a circle. It presupposes the truth of that very thing that is under debate, namely, that the government has a right to kick people out of "its" territory if they do not comply with its onerous regulations. If we jettison this basic unproven premise, then there is no case for inviting people to "leave" who object to statist depredations. Another difficulty is that there were people living on the territory taken over by the government before it was set up. This is necessarily so. There had to be people in existence to set up a government before it was operational. Thus, these people were there *before* the advent of the government. They could with as much justification say to those in the process of forming a government: "If you don't like it here, under our present state of free market anarchism, you are free to leave." Why must those who object to this obnoxious institution be the ones who must leave, on logical grounds?

Let us attempt to further "hack away" at this truly evil type of organization. It is said that the reason we need a government is because two individuals, or groups of people, will fight each other without such an institution, and, if allowed, this would create chaos. Or bullies would prey on

the weak. There are great difficulties with this view. First of all, as a matter of historical fact (Rummel 1994), governments have murdered more than 150 millions of their own citizens in the last century, and this is apart from the wars they have fomented with each other. And this statistic does not even include some 40,000 people slaughtered on US government highways (Block 2009). Gangsters and bullies are simply not in the same category. Secondly, a logical implication of national government is world government. If we need a national government because people in that territory will fight with each other, then we need a world government since national governments will (and most certainly have) fight with each other. The problem with world government is of course that there will be nowhere to run when and if such an institution turns totalitarian. But if world government is rejected, so, then, must the same fate await national government, since there is a perfect analogy between the two cases.

Here is yet another attempt to justify the unjustifiable: the government gives us goods and services in return for the taxation it imposes on us. This, too, cannot be denied. The state provides a myriad of goods and services, ranging from battleships to postal service, from welfare to food stamps, from bailouts to Detroit, to Wall Street, and to numerous points in between. But suppose a philosophical mugger who is willing to engage in discourse were to hold up an innocent man. The latter objects on the ground that the robber is not providing any goods or services, whereupon the thief agrees with his victim, amends his evil ways, and offers him the services of a smile or a snarl, and, also, a paper clip or a rubber band. The point is, it is entirely irrelevant whether or not the private or governmental burglar offers his prey anything in return for his depredations or not. The only relevant issue is if the citizen did *agree* to the interaction or not, and, as we have seen, not only is there no evidence that everyone unanimously agreed, but also there *could not* have been any such occurrence. For, if it did, if, that is, all the citizens welcomed the government, that would be a logical contradiction. For government is defined as that institution about which everyone most patently *did not* agree. If there were an organization to which all participants voluntarily subscribed, it would no longer be a government! Rather, it would be some sort of private institution. A very large one, but similar in all essential regards to a business firm. For the latter embodies "capitalists acts between consenting adults" in the felicitous phraseology of Nozick (1974, p. 163).

Cross-References

▶ Organization

References

Block WE (2009) The privatization of roads and highways: human and economic factors. The Mises Institute, Auburn. http://www.amazon.com/Privatization-Roads-And-Highways-Factors/dp/1279887303/ref=sr_1_1?s=books&ie=UTF8&qid=1336605800&sr=1-1

Nozick R (1974) Anarchy, state and utopia. Basic Books, New York

Rothbard MN (1977) Do you hate the state?. In: The libertarian forum, Joseph Peden, New York vol 10, no 7. http://www.lewrockwell.com/rothbard/rothbard75.html

Rummel RJ (1994) Death by government. Transaction, New Brunswick. http://www.hawaii.edu/powerkills/NOTE1.HTM; http://www.hawaii.edu/powerkills/20TH.HTM

Schumpeter JA (1942) Capitalism, socialism and democracy. Harper, New York

Spooner L (1966[1870]) No treason: the constitution of no authority and a letter to Thomas F. Bayard. Rampart College, Larkspur. http://jim.com/treason.htm

Government Failure

Giuseppe Di Vita
Department of Economics and Business, University of Catania, Catania, Italy

Abstract

The analysis begins from the origin of the concept of market failure and continues with an explanation of the reasons why a

measure of political economy could worsen market allocation and the level of social welfare. Mention is made of the differences between Europe and the United States regarding the weight of the State in the economy and of the efforts made by research to give a quantitative measure to government failure.

Definition

Government failure occurs when a measure of economic policy or the inactivity of the government worsens the market allocation of resources reducing economic welfare.

Government Failure

From a reading of the first few chapters of textbooks of *Principles of Economics*, we have learned that under the assumption of a perfect competitive market, for each good (or bad) exchanged, the economy achieves its maximum level of welfare as a result of the Smithian theory of the "invisible hand" (Smith 1776). This means that in such a Pareto-optimal situation, there is no room for economic policy measures that, if adopted, may only constitute an obstacle to the proper functioning of the market.

Despite this unlimited confidence in the market virtues, the same Adam Smith was forced to recognize the limits of the market, regarding national defense, justice, and public services (i.e., waterworks, hospitals, schools, roads, etc.), for which there is no profit in producing at an individual level, because the costs are always greater than the revenues (Smith 1776). Although one of the fathers of the modern economic theory recognized the limits of the market, the dominant view, up until the 1940s, was that the government should not interfere with the market to modify or change market allocation or income distribution.

The Great Depression started in the late 1930s and the subsequent Keynesian theory (Keynes 1936) was further instrumental in destroying the myth of the free market. After the Second World War, the economic theory started to consider than before the existence and causes of market failure more seriously (Bator 1958) and saw the need for some government intervention, the so-called visible hand (Chandler 1977), to improve market allocation and raise the level of social welfare.

During the second half of the last century, the broad consensus regarding the positive consequences of adopting economic policy to correct the effects of market failures on economic welfare Coase (1964) introduced into the economic literature the expression "government failure," as opposed to "market failure," to express some concern about the continual assumption of Pareto-improving consequences of economic policy measures (Posner 1969, 1974). (Ronald Coase (1964) agreed with the approach which evaluated the effects of economic policy measures comparing regulated industries with industries not subject to regulation, although this approach cannot be followed because it is difficult to find industries that are comparable regardless of the degree of regulation.)

Coase drew inspiration from the talk of Professor Roger C. Cramton, a law scholar, held in Boston during the Seventy-Sixth Annual Meeting of the American Economic Association: Cramton stated that lawyers "...focus on the fact that public officials and tribunals are going to be fallible at best and incompetent or abusive at worst...." Coase (1964) remarks that this statement is completely opposed to the assumption that the economists make about the government. In the economic theory, the government is seen as a benevolent planner who wants to correct the defects of the invisible hand, to improve market allocation, and to raise the level of welfare. The opinion of the government expressed by Professor Crampton would probably have been considered provocative, but for the economist, it was an alarm bell underlining the necessity to rethink the assumption about the government bodies and the impact of economic policy measures on the level of welfare. (In general, about the measure of economic policy, we can report the words of Milton Friedman (1975): "...The government solution to a problem is usually as bad as the problem and very often makes the problem worse....")

G

Moreover, Coase (1964) emphasizes the inability of the government to supply an immediate answer to the changes in economic conditions and the crucial role of "detailed knowledge" (or information) of the economic phenomenon addressed by a measure of economic policy, as an indefectible condition of government intervention. (Winston (2006) has recently affirmed that government failure "...arises when government has created inefficiencies because it should not have intervened [in the market] in the first place, or when it could have solved a given problem or set of problems more efficiently, that is, by generating greater net benefits..."; such a definition empathizes the inability of the government to improve market allocation.)

It is possible to affirm that a government policy is worth adopting in the presence of a market failure that is a source of not negligible social costs (see Datta-Chaudhuri (1990), Krueger (1990), Wallis and Dollery (1999, 2001), and Wiesner (1998) about the policy measures adopted in developing countries, since the 1960s, as sources of economic inefficiency which worsen the welfare level) where the government policy is at least improving market performance and efficiently correcting the market failure and optimizing the economic welfare (for normative economic theories of government failure, see Dollery and Worthington 1996).

The first possible source of government failure is the dynamic nature of an economic process such that the government in charge cannot commit itself for future measures which will be adopted by the subsequent government. (The dynamic nature of the economic process makes it necessary to consider the discount rate in the analysis: this is a task that we leave to future and more detailed analysis.) Stiglitz (1998) offers the example of hydroelectricity plants to produce electricity which need to be subsidized for a period of time longer than the duration of the legislature, while the hydroelectricity industries have no guarantee that successive governments will leave unaltered the subsidies and the legislation. (The inability to make commitments causes another set of inefficiencies: the cost of creating next-best credibility-enhancing mechanisms. While those in the government at one date cannot commit future governments, they can affect the transaction costs of reversing policies (Stiglitz 1998, p. 10).)

The economic policy measures worsen market allocation whenever there is imperfect information between the policy maker and the private parties involved. Imperfect information may represent an obstacle to improving Pareto efficiency, if the policy maker does not have access to all the relevant information to adopt the correct policy measure or establish the magnitude of the measure (e.g., the amount of the rate of taxation).

The third source of government failure is the so-called destruction competition, a process that occurs in the presence of imperfect competition. Firms can get ahead not just by producing a better product at lower costs but also by raising the costs of their rivals (Salop and Scheffman 1983). Destructive competition is a zero-sum game where the profit of one is at the expense of another (see, e.g., the political games, with positions to be won or lost). This means that under the imperfect competition market structure, the liberalization of the market or the inactivity of the policy maker (Orbach 2013), as a result of a rational choice, determines the achievement of a suboptimal level of social welfare.

Another source of government failure is uncertainty about the consequences of policies: this may be due, as in the case of economic policy measures taken under asymmetric information, to the inability to predict the future impact of economic policies adopted now. A lesson we learned is that to be effective, economic policy measures should have well-defined objectives to achieve; otherwise, the measures taken by the government without a clear and precise statement of the objectives to be pursued and of the resources to meet them are just a way to leave unchanged the allocation of resources and the distribution of income (Cramton 1964). This means that some cases that may be attributable to the concept of government failure are just from the point of view of social welfare, because the government policies worsen market allocation but satisfy the real will of the policy maker to protect some established interests, without displeasing public opinion.

Finally, the last point we address is that government policies may be constrained in their action and worsen welfare due to the complexity of the economic measures that are hard for public opinion to understand, so Stiglitz (1998) empathizes the "simplicity constraint" in economic policies. Simple policies are easily explained to, and approved by, public opinion.

In a more interdisciplinary environment, considering laws, institutions, and economics, the theory of government failure has recently been enriched, affirming that it may also be due to the inefficient designing of rules for the economy that may be excessively specific (i.e., standard), too broad, conflicting, and unfair (Dolfsma 2011). (In general, on inefficient regulation as a source of government failure, see Posner (1974).)

The theories of government failure have played a different role in the United States and in the Old Continent because in the former they have been used to justify the limited adoption of economic policy measures and a reduced role of the state within the economy. In Europe, the limits of government action have been used to justify a reduction of the weight of the public sector in the economy (Vickers and Yarrow 1991).

Despite the differences between Europe and the United States in the role of the state in the economy that can be measured by the tax burden (Romer and Romer 2010), the government follows the business cycle in its policies because during crises, it is forced by public opinion to adopt stronger measures to curb the crisis (Rajan 2009), while during periods of prosperity, it is more prone to pander to the desire of firms for freedom.

Winston (2006), basing his theories on an empirical research limited to antitrust, monopsony policy, and economic regulation to curb market power, so-called social regulatory policies to correct imperfect information and externalities, and public production to provide socially desirable services, reached the conclusion that government intervention in markets has either been unnecessary or has missed significant opportunities to improve performance. (For recent development of the theory of "government failure," see Mitchell and Simmons (1994) and Winston (2006).)

The economic policy measures to correct the limitations of markets have played a significant role in economic history (Datta-Chaudhuri 1990) and will probably continue to be useful in the future, but to improve market allocation, the policy maker should be aware of the limits of the "visible hand" (Winston 2012).

Cross-References

▶ Government
▶ Market Failure: Analysis
▶ Market Failure: History
▶ Public Goods

G

References

Bator FM (1958) The anatomy of market failure. Q J Econ 72(3):351–379

Chandler A (1977) The visible hand. Belknap, Cambridge, MA

Coase R (1964) The regulated industries: discussion. Am Econ Rev 54(2):194–197

Cramton RC (1964) The effectiveness of economic regulation: a legal view. Am Econ Rev 54(3):l82–l191

Datta-Chaudhuri M (1990) Market failure and government failure. J Econ Perspect 4(3):25–39

Dolfsma W (2011) Government failure -four types. J Econ Issues XLV(3):593–604

Dollery B, Worthington A (1996) The evaluation of public policy: normative economic theories of government failure. J Interdiscip Econ 7(1):27–39

Friedman M (1975) An economist's protest, 2nd edn. Thomas Horton & Daughters, Sun Lakes

Keynes JM (1936) The general theory of employment, interest and money. Macmillan, London

Krueger AO (1990) Government failure in development. J Econ Perspect 4(3):9–23

Mitchell WC, Simmons RT (1994) Beyond politics: markets, welfare, and the failure of bureaucracy. Westview Press, Boulder

Orbach B (2013) What is government failure? Yale J Regul 30:44–56

Posner RA (1969) Natural monopoly and its regulation. Stanf Law Rev 21:548–643

Posner RA (1974) Theories of economic regulation. Bell J Econ 5(2):335–358

Rajan R (2009) The credit crisis and cycle-proof regulation. Fed Res Bank St Louis Rev 91(5): 397–402

Romer CD, Romer DH (2010) The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. Am Econ Rev 100:763–801

Salop SC, Scheffman DT (1983) Raising rivals' costs. Am Econ Rev 73(2):267–271
Smith A (1776) An inquiry into the nature and causes of the wealth of nations. W. Strahan and T Cadell, London
Stiglitz JE (1998) Distinguished lecture on economics in government: the private uses of public interests: incentives and institutions. J Econ Perspect 12:3–22
Vickers J, Yarrow G (1991) Economic perspectives on privatization. J Econ Perspect 5(2):111–132
Wallis J, Dollery B (1999) Market failure, government failure, leadership and public policy. St. Martin's Press, New York
Wallis J, Dollery B (2001) Government failure, social capital and the appropriateness of the New Zealand model for public sector reform in developing countries. World Dev 29(2):245–263
Wiesner E (1998) Transaction cost economics and public sector rent-seeking in developing countries: toward a theory of government failure. In: Picciotto R, Durán EW (eds) Evaluation and development: the institutional dimension, World Bank Editions. Transaction Publishers, Washington, DC
Winston C (2006) Government failure versus market failure. Microeconomics policy research and government performance. R. R Donnelley , Harrisonburg
Winston C (2012) Government policy for a partially deregulated industry: deregulate it fully. Am Econ Rev 102(3):391–395

Government Quality

Andreas P. Kyriacou
Universitat de Girona, Girona, Spain

Abstract

A growing body of work in economics has pointed towards the crucial importance of government quality for economic development. A major issue emerging in related empirical work has been the need to account for the confounding influence of other factors as well as the presence of reverse causality or the likelihood that development itself may facilitate governance. The potentially key role of government quality in explaining international differences in economic development has led many scholars to try and identify those factors that determine it. This research effort has brought forth the role of numerous variables including social heterogeneity, cultural heritage, and constitutional design. The lion's share of empirical work has employed measures of government quality based on perceptions. A debate exists about the convenience of conflating government quality with institutional quality. To measure institutional quality, it may be more advisable to generate measures or indicators which capture the extent to which institutions actually constrain governments.

Definition

Government quality is a multidimensional concept. Government quality is said to improve if the public sector does not distort the proper functioning of the private sector (respects property rights, does not overregulate), is an efficient administrator (low corruption, less bureaucracy, high tax compliance), provides public goods (education, health, infrastructures, etc.), and, finally, allows for and protects political freedom.

Government Quality

Over the years, economists have come to identify government or institutional quality as a fundamental cause of economic development. Early on, North and Thomas (1973) argued that the establishment of private property rights was instrumental for the rise of the West since the middle ages. Inspired by this insight, a great deal of scholarship has empirically explored the link between institutional quality and development (most notably Hall and Jones 1999; Acemoglu et al. 2001; 2002; Rodrik et al. 2004).

Growth-promoting institutions have at least two features: they must enforce property rights for a broad cross section of society so that all individuals have an incentive to invest, innovate, and take part in economic activity. And they must furnish some degree of equality of opportunity in society, including equality before the law (so that those with good investment opportunities can take advantage of them), and equal access to health and education (to promote human capital and to account for the fact that entrepreneurial initiative

may be randomly distributed in the population and as such be independent of the distribution of property rights to existing resources) (Acemoglu et al. 2005).

A major concern which emerges when attempting to calibrate the impact of government quality on development is the issue of reverse causality or, in other words, the expectation that development facilitates institutional quality. In this vein, it has been argued that economic development makes better quality institutions more affordable (Islam and Montenegro 2002) and will tend to create a demand for better government (La Porta et al. 1999), perhaps because of income's positive effect on education, literacy, and depersonalized relationships (Treisman 2000). To deal with reverse causality, scholars have employed various instruments of government quality. Arguably the most notable has been settler mortality in the New World proposed by Acemoglu et al (2001, 2002). They argue that in those colonies with high indigenous population densities where the disease environment was inimical to mass European settlement, extractive institutions, with few limits on the capacity of the state to expropriate individuals, were set up. And these institutions persist over time, not least because of the interest of the extractive elite to maintain their privileges. Conversely, in those colonies which were sparsely populated where, moreover, the disease environment was more benign, mass European settlement occurred leading to demands for good quality government (c.f. Sokoloff and Engerman 2000). Cross-country empirical work based on settler mortality suggests that government quality is a key factor explaining international differences in economic development, even trumping the direct effect of geography and trade (Rodrik et al. 2004).

The robustness of these finding has not gone unchallenged. On the one hand, the quality of the mortality data has been the subject of some debate among scholars (Albouy 2012; Acemoglu et al. 2012). On the other, it has been argued that the pattern of European settlement may have influenced growth through other, more fundamental, channels, namely, human capital. Thus, European colonizers took with them their human capital, and in those colonies which were more extensively settled, these human capital endowments enhanced both institutional and productive capacities leading to stronger economic development (Glaeser et al. 2004).

Notwithstanding these qualifications, and based on theoretical and empirical work pointing towards a key role of institutions for the wealth of nations, scholars have tried to identify those factors that determine institutional quality, beyond the level of economic development itself and initial conditions framed by climate and factor endowments. In particular it has been variously argued that government quality may be determined by a range of variables which include interpersonal income inequalities (You and Khagram 2005; Glaeser et al. 2003; Sonin 2003), ethnic heterogeneity (Alesina et al. 1999; Glaeser and Saks 2006; Alesina and Zhuravskaya 2011), and ethnic group inequalities (Baldwin and Huber 2010; Kyriacou 2013). In addition, international differences in government quality may be due to cross-country cultural differences (Fisman and Miguel 2007; Licht et al. 2007) which may be partly the result of different religious heritages (La Porta et al. 1999; North et al. 2013). Government quality may also be responsive to constitutional design including electoral rules and government systems (Persson et al. 2003; Kunicová and Rose-Ackerman 2005) and the degree of fiscal and political decentralization (Fisman and Gatti 2002; Enikolopov and Zhuravskaya 2007).

Quantitative work in law, economics, political science, and business administration has employed government quality indicators from a range of sources. These include the World Bank's Worldwide Governance Indicators, Transparency International's Corruption Perception Index, the World Economic Forum's Executive Opinion Survey, the International Country Risk Guide (ICRG) published by the Political Risk Services Group, and several indicators elaborated by Freedom House and the Fraser Institute. Because of their country coverage and comparability over time, many experts have relied on indicators from the ICRG – available since 1984 – as well as the Worldwide Governance

Indicators, available since 1996. The ICRG data are based on in-house expert assessments of different dimensions of government quality including government stability, contract viability or expropriation, corruption, law and order, democratic accountability, and bureaucratic quality.

Unlike the ICRG indicators which are derived from a single source, the Worldwide Governance Indicators are obtained from a host of perception-based sources, including the ICRG and the other sources listed above. The Worldwide Governance Indicators measure government quality across six dimensions: (1) voice and accountability, (2) political stability and the absence of violence, (3) government effectiveness, (4) regulatory quality, (5) the rule of law, and (6) control of corruption. A key feature of the indicators is that all country scores are accompanied by standard errors which reflect the number of sources available for a country and the extent to which these sources agree with each other (Kaufmann et al. 2010). It has been pointed out that the indicators measuring the last four dimensions are very strongly correlated and, as such, they may be measuring the same thing (Langbein and Knack 2010). As a result, it makes sense to combine these four into an aggregate indicator in empirical work. Moreover, it may be useful to view the first two dimensions as contributing towards the latter four.

There is some debate about conflating institutional quality with government quality. According to Douglass North (1991, p. 97) “institutions are the humanly devised constraints that structure political, economic and social interaction.” From this perspective, institutions are constitutions and other laws, and the test of good institutions would imply the examination of how such constraints improve social welfare. Because empirical work has tended to equate institutional quality with government quality, scholars have measured the former by way of the perception-based indicators reviewed above. But government quality indicators typically measure outcomes rather than institutional constraints per se, and objective measures of institutional constraints are typically uncorrelated with measures of government quality (Glaeser et al. 2004). However, empirical efforts relating de jure institutional constraints with social outcomes are likely to be undermined by the possibility that the measures or indicators used may not capture the extent to which constraints are actually binding (Voigt 2013).

References

Acemoglu D, Johnson S, Robinson J (2001) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91:1369–1401

Acemoglu D, Johnson S, Robinson J (2002) Reversal of fortune: geography and institutions in the making of the modern world income distribution. Q J Econ 117:1231–1294

Acemoglu D, Johnson S, Robinson J (2005) Institutions as a fundamental cause of long-run growth. In: Aghion P, Durlauf S (eds) Handbook of economic growth. Elsevier, Amsterdam, pp 385–472

Acemoglu D, Johnson S, Robinson J (2012) The colonial origins of comparative development: an empirical investigation: reply. Am Econ Rev 102(6): 3077–3110

Albouy D (2012) The colonial origins of comparative development: an empirical investigation: comment. Am Econ Rev 102(6):3059–3076

Alesina A, Zhuravskaya E (2011) Segregation and the quality of government in a cross-section of countries. Am Econ Rev 101:1872–1911

Alesina A, Baqir R, Easterly W (1999) Public goods and ethnic divisions. Q J Econ 114(4):1243–1284

Baldwin K, Huber J (2010) Economic versus cultural differences: forms of ethnic diversity and public good provision. Am Polit Sci Rev 104(4):644–662

Enikolopov R, Zhuravskaya E (2007) Decentralization and political institutions. J Public Econ 91:2261–2290

Fisman R, Gatti R (2002) Decentralization and corruption: evidence across countries. J Public Econ 83: 325–345

Fisman R, Miguel E (2007) Corruption, norms and legal enforcement. Evidence from diplomatic parking tickets. J Polit Econ 115(6):1020–1048

Glaeser E, Saks R (2006) Corruption in America. J Public Econ 90:1053–1072

Glaeser E, Scheinkman J, Shleifer A (2003) The injustice of inequality. J Monet Econ 50:199–222

Glaeser E, La Porta R, Lopez-de-Silanes F, Shleifer A (2004) Do institutions cause growth? J Econ Growth 9:271–303

Hall R, Jones C (1999) Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114(1):83–116

Islam R, Montenegro C (2002) What determines the quality of institutions? Background paper for the World Development Report 2002, Building Institutions for Markets, Washington DC

Kaufmann D, Kraay A, Mastrutzzi M (2010) The worldwide governance indicators: methodology and

analytical issues. World Bank Policy research paper no 5430, Washington DC
Kunicová J, Rose-Ackerman S (2005) Electoral rules and constitutional structures as constraints on corruption. Br J Polit Sci 35(4):573–606
Kyriacou A (2013) Ethnic group inequalities and governance: evidence from developing countries. Kyklos 66(1):78–101
La Porta R, Lopez-de-Silanes F, Shleifer A, Vishny R (1999) The quality of government. J Law Econ Organ 15(1):222–279
Langbein L, Knack S (2010) The worldwide World Governance Indicators: Six, one, or none? J Dev Stud 46(2):350–370
Licht A, Goldsmith C, Schwartz S (2007) Culture rules: the foundations of the rule of law and other norms of governance. J Comp Econ 35:659–688
North D (1991) Institutions. J Econ Perspect 5:97–112
North D, Thomas R (1973) The rise of the western world. A new economic history. W.W. Norton, New York
North C, Orman W, Gwin C (2013) Religion, corruption, and the rule of law. J Money Credit Bank 45(5): 757–779
Persson T, Tabellini G, Trebbi F (2003) Electoral rules and corruption. J Eur Econ Assoc 1(4):958–989
Rodrik D, Subramanian A, Trebbi F (2004) Institutions rule: the primacy of institutions over geography and integration in economic development. J Econ Growth 9:131–165
Sokoloff K, Engerman S (2000) History lessons: institutions, factor endowments, and paths of development in the new world. J Econ Perspect 14:217–232
Sonin K (2003) Why the rich may prefer poor protection of property rights. J Comp Econ 31:715–731
Treisman D (2000) The causes of corruption: a cross-national study. J Public Econ 76:399–457
Voigt S (2013) How (not) to measure institutions. J Inst Econ 9(1):1–26
You J, Khagram S (2005) A comparative study of inequality and corruption. Am Sociol Rev 70:136–157

Greece: Ancient Greece

Michel S. Zouboulakis
Department of Economics, University of Thessaly, Volos, Greece

Abstract

The Ancient Greek miracle is the synthesis of institutionally constrained free citizens seeking for their personal improvement within a democratically evolving law structure which guarantees public interest and social justice.

Definition

What we call ancient Greece was actually a large number of dispersed and independent city-states with a common language, religion, and social customs and yet with quite distinct political, social, and economic institutions. Chronologically, what we call Ancient Greek world extends from the eighth century BC to the second century AD. Geographically, the Ancient Greek world covered parts of Southern Italy and Sicily in the west, as well as the entire coast on the other side of the Aegean and the Black Sea, and many other parts of the Mediterranean, including North Africa, Southern France, and Spain. Our purpose here is to focus on a restricted part of Ancient Greek social and economic institutions, saying the minimum about the political institutions and nothing about the informal ones, meaning the values and social norms which shaped this glorious era of the western civilization.

Population and Citizenship

Cities were organized around a well-defined geographical center, and their population size varied from 250,000 in fifth-century Athens to 3,000 in the city-state of the island of Aegina. A common characteristic among all Greek cities was that not all the inhabitants were recognized as citizens with full political rights and civil duties. In general, citizens were free males above the age of 18, having completed their military duties. But not all adult free citizens, even if they were born in the city, had full citizenship. Noncitizens who were born free or were freed in their lifetime were often called and had military and tax obligations. An important part of every Greek city consisted of slaves, in an analogy (very different from city to city) one citizen for every two or three slaves. Social stratification differed from one city-state to another, although there were many common characteristics. In Athens, Solon introduced his institutional reforms in 594 BC, based on the division of citizens according to their income. Citizens, and therefore their obligations toward their city, were defined according to their

production: the *pentakosiomedimnoi* producing more than 500 *medimnos* (approximately 260,000 lit.); the *ippeis,* producing more than 300 *medimnos* or 156,000 lit.; the *zeugitai*, with more than 200 *medimnos* or 104,000 lit.; and finally the *thetes*, with less than 200 per year. Among these four property classes, the poorest had limited public rights; the *thetes* could vote in the Assembly but were excluded from public office. The richer classes provided the main military effort, as cavaliers (*ippeis*) or as heavily armed foot soldiers (*hoplites*). The situation in Sparta was different. Descendents of Spartan citizens got full citizenship at the age of 18, having received a proper military education called *agoge*. Other inhabitants were the *perioikoi*, who were free to live in Spartan territory but were noncitizens, and the *helots*, the state-owned serfs (Glotz 1928; Finley 1964; Cartledge 1993; Sakellariou 1999).

Political Institutions

Politically, the majority of Greek cities were ruled by aristocracies, oligarchies, and tyrannies in the eighth, seventh, and sixth centuries and oligarchies and democracies in the fifth and fourth. In Corinth, Argos, Sparta, Thebes, and many other cities, oligarchy was based on property. Aristotle described four types of oligarchy from the most autocratic cities of Thessaly in central Greece to the more open cities of Thebes and Orchomenos in Boeotia; participation in the privileged ruling few (*oligoi*) was only a matter of middle-range personal revenue among full citizens. Oligarchic cities such as Croton, Region, Acragas, and Colophon had very large Assemblies (*ecclesiae*) consisting of 1,000 wealthy citizens who legislated by open vote. Sparta had a smaller "executive" body of 30 elders (*gerousia*) advising the hereditary two kings (Glotz 1928; Cartledge 1993).

Athenian direct democracy of the fifth century was an exception that followed well-established formal rules and had two collective bodies of governance: the *Boulé,* a body of 500 citizens selected by draw for a year term from the ten counties of Attica, met every day, except on holy festivals, to administrate the city. All the citizens (i.e., all 42,000 in 431 BC) had the right to participate in the general assembly (*Ecclesiae*) 10–40 times in a year at the *Pnyx*, a hill opposite to the Acropolis, to decide freely on all the important public matters of the city and to legislate. The farther one lived, the less often he was present to vote. So, distant communes of Attica tended to be underrepresented, this being one of the main problems of Athenian democracy, together with the lack of material motives to participate. Poor citizens would be less willing to participate in the Ecclesiae, not affording to lose one day's wages. After the reforms of Pericles in 460 BC, citizens were paid a small remuneration, little less than a day's wage for their participation in the Assembly (Glotz 1928; Mossé 1995). But, there were clear advantages as well. Since the voting rule was simple majority by show of hands, it was quite difficult for organized groups to carry the day if they could not convince the majority of the citizens, many of whom changed their minds after hearing the orators' arguments. There were no formal political parties in the modern sense, but there certainly were political groups representing interests, gathered around some charismatic leaders, such as Cleisthenes, Themistocles, Ephialtes, and Pericles for the democratic "party" and Aristides, Kimon, and Nikias for the aristocratic or rather conservative "party." The latter was mainly supported by the wealthier landowners and also by medium holders whose revenues came from agriculture, while the farmers and breeders, as well as the traders, artisans, and sailors, supported the democratic side (Finley 1964).

Private and Public Finance and Institutionalized Benevolence

In the great majority of archaic and classical Greek, city-state inhabitants lived mainly of agriculture. Even the rich Athens, Corinth, and Chalkis were mainly inward-looking city-states, but they all had a significant export trade and a commercial fleet as well. Archeological finds in

Italy and Asia Minor prove the flourishing trade of other Greek naval city-states such as Miletus, Ephesus, Colophon, Sybaris, Syracuse, Acragas, and even Massalia (Marseilles). Since the seventh century, golden, silver, and electron coins contributed to the rapid monetization of trade and economic transactions in general. Economic development rapidly changed the life of these naval city-states, while mainland agricultural cities in Boeotia and the Peloponnese declined slowly in power and wealth (Glotz 1928). Individual banks (*trapezes*) appeared early enough to assist commerce with credit and change facilities. Although the institution of the commercial bank appeared only in the Renaissance, ancient capital owners actually acted as bankers lending money with interest to farmers and naval traders (Bitros and Karayiannis 2010). The fact that bankers differentiated interest rates according to the purpose of the loan, offering loans with a nominal rate of 12 % to landowners and up to 30 % to shipowners, proves that they had a clear notion of risk.

Classical Athens had a highly developed system of property rights and duties. Despite property-based distinctions of social rank, inherent to all hierarchical societies, tax burdens and public and military obligations were specified according to each one's income situation. Hence, the city-state of Athens was able to establish an exact contract with its richest citizens in order to ensure the production of both private and public goods and to secure a permanent source of state revenue. The city's public revenues came from rents of public land, custom duties, fines, and war booty. In many Greek cities, rich citizens contributed in a semi-voluntary basis to the public expenditure according to the social institution of *liturgy*. The most important liturgies were the *choregia*, the organization and finance of theatrical productions; the *gymnasiarchia*, the financial support and supervision of athletic training; and the *estiasis*, the organization of public religious feasts (Davies 1981; Gabrielsen 1994; Mossé 1995). Hence, when Athens decided to build 100 ships by means of the silver discovered at Lavrion, to defend the city from the second Persian invasion, an institutional reform introduced a new liturgy, called *trierarchy*. It consisted of entrusting to the 100 richest citizens the charge of building 100 more new ships so that the Athenian fleet was comprised of 200 triremes, equivalent to the two thirds of the fleet that defeated the Persians in the naval battle of Salamis in 480 BC (Kyriazis and Zouboulakis 2004).

After Salamis, the *trierarchy* system was further developed and refined. Under its new form, the city itself financed the construction of warships but entrusted their maintenance to rich citizens, elected each year. Eligibility was based on the formal rule of property qualification (Glotz 1928; Davies 1981). If an Athenian citizen was able to prove before the Court that there was a richer citizen than himself, then he could avoid the duty. If the supposedly richer man disagreed with the terms, he had to choose to refuse the obligation and exchange properties with the challenger or to undertake the liturgy. This procedure was called *antidosis* (Gabrielsen 1994; Carmichael 1997). A refusal to discharge any liturgy entailed punishment. While the city was safeguarded against fraud by the liturgists by mortgaging their land assets and houses, rich Athenians considered it a duty and an honor to be chosen to finance and manage the production of public goods and services (Lyttkens 1997). To avoid that liturgies do not fall on the same persons for consecutive years, after 378 BC taxpayers were organized in 20 groups called *symmoriae* to share the burden (Mossé 1995).

The Athenian tax system was comprised also of market dues, a tax on sales concluded before state officials, an annual charge of 12 drachmae on *metics* and of 0.5 drachmae on slaves and freedmen, a tax on those who pursued certain callings requiring special supervision (such as oracle mongers, jugglers, and prostitutes), as well as import and export duties. Direct taxation levied on land during the sixth century BC was abolished because it was regarded as an infringement on civil liberty except in times of crises, like the Peloponnesian War, when the Athenians imposed upon themselves a special property war tax *eisphora*. Still, the liturgies functioned as a kind of progressive direct wealth taxation. Redistribution followed indirectly, since the taking over of public goods by the richest citizens liberated

public means to be spent on other purposes, such as a "poor-relief" system, granting assistance to the incapacitated and maintaining the children of those who fell at war, until they came of age.

As it was argued (Ackroyd 1992), lessons from Classical Greece demonstrate that private interests are not incompatible with social justice and public interest insofar as these values are not imposed by a central public power but are enforced by the Law decided by the citizens themselves (ΝΟΜΟΣ ΒΑΣΙΛΕΥΣ).

Cross-References

- ► Constitutional Evolution in Ancient Athens
- ► Development and Property Rights
- ► Fiscal System
- ► Institution
- ► Public Enforcement

References

Ackroyd P (1992) Greek lessons for property right arrangements: justice and nature protection. Am J Econ Sociol 51:19–26

Bitros G, Karayiannis AD (2010) Morality, institutions and the wealth of nations. Some lessons from Ancient Greece. Eur J Polit Econ 26:68–81

Carmichael C (1997) Public munificence for private benefit: liturgies in Classical Athens. Econ Inq 35: 261–270

Cartledge P (1993) The Greeks. Oxford University Press, Oxford

Davies JK (1981) Wealth and the power of wealth in Classical Athens. Arno Press, New York

Finley M (1964) The Ancient Greeks. La Découverte, Paris, (French translation), 1984

Gabrielsen V (1994) Financing the Athenian fleet: public taxation and social relations. Johns Hopkins University Press, Baltimore

Glotz G (1928) La cité grecque. Albin Michel, Paris, 1953

Kyriazis N, Zouboulakis M (2004) Democracy, Sea power and institutional change: an economic analysis of the Athenian Naval Law. Eur J Law Econ 17: 117–132

Lyttkens CH (1997) A rational actor perspective on the origins of liturgies in Ancient Greece. J Inst Theor Econ 153:462–484

Mossé C (1995) Politique et société en Grèce ancienne. Le "modèle" athénien. Aubier, Paris

Sakellariou MB (1999) The Athenian democracy. University of Crete Press, Heraklion

Greece: Modern Greece 1821–2018, A political History of

Aristides Hatzis
Department of History and Philosophy of Science, National and Kapodistrian University of Athens, Athens, Greece

Definition

Modern Greece has a history of almost two centuries. During these centuries, the country managed to move from the backwaters of Europe to a prosperous liberal democracy before economic crisis hit the country hard in 2010. Greece was founded after a War of Independence from the Ottoman Empire that was based on liberal and democratic principles. This left a political legacy which led to universal male suffrage as early as 1844 and one of the longest parliamentary histories in Europe, despite the tumultuous political life and brief periods of authoritarian regimes. The nineteenth century was a period of a slow modernization of the country (in infrastructure and institutions) but is was also suffocated by "Megali Idea," the irredentist dream of the enlargement of the Greek state to include all lands, under Ottoman rule, inhabited by large Greek-speaking populations. A great part of Megali Idea was realized in early twentieth century but the triumphs ended with a devastating catastrophe in 1922. Greek political elites were often incompetent and corrupt, but several reformist statesmen managed gradually to achieve convergence with other western European countries. Most importantly, they were very effective in steering Greece on the right (i.e., winning) side of history during every major European or Global conflict (Balkan Wars, World Wars, Cold War). Greece, after World War II and a ferocious Civil War, enjoyed one of the strongest, almost uninterrupted growth on a global level. This led to the accession to the European Communities in 1981 and later the Eurozone. Today, after 10 years of economic crisis and painful austerity,

Greece must meet one of the most difficult challenges: to achieve growth by adopting inclusive institutions.

Greeks in the Ottoman Empire

Greece became an independent state in 1830. Its independence was the result of a national uprising against the Ottoman Empire in the early nineteenth century. After the end of the classical era, Greece was controlled by empires, mostly by the Roman, the Eastern Roman or Byzantine and the Ottoman empires. The Roman and the Byzantine empires were strongly influenced by the ancient Greek civilization and from the seventh century on, Byzantine Empire was linguistically Hellenized.

After the fall of Constantinople, in 1453, the Ottoman Empire recognized the Christian Orthodox Church and the Patriarch of Constantinople as the spiritual but also the political leader of the Greek Christian Orthodox (*Rum*) community. The Christian Orthodox Church, which was dominated by a Greek speaking clergy, was granted several privileges, including some autonomy and judicial jurisdiction in certain private law disputes among Orthodox Christians.

As a result of the power and the great influence of the Ecumenical Patriarchate of Constantinople, at the beginning of the nineteenth century, the Greek Christian Orthodox Commonwealth was so prominent as to capture even high-ranking political stations in the Ottoman administration. Since the late seventeenth century, Greek Christian Orthodox laypeople, "Phanariotes," who inhabited Phanare, the area around the seat of the Patriarch, had a significant political power. Phanariotes dominated the influential position of Grand Dragoman, the official interpreter for the Imperial Council, but also the position of the Prince (*Voivode*) of the semi-autonomous Danubian Principalities of Moldavia and Wallachia, i.e., modern-day Romania. Phanariotes, many rich merchants, several local communities, and the Church created, from seventeenth century on, a big number of local but also influential small schools which became, until the nineteenth century, the standard educational units in the Balkan Christian Orthodox communities.

There was also a local Christian Orthodox landed class in Southern Greece, mostly in Peloponnese. The *Kodjabashis* in Turkish or *Proestoi* in Greek were the leaders of their communities. They were part of the Ottoman administration since they were usually responsible for tax collection. These local notables were supported by private armed groups, and they had direct connections to the Ottoman administration and the Patriarchate. They had to deal with dangerous bands of brigands (*klephts*) who had managed to control the mountainous areas of Peloponnese and Central Greece (*Roumeli*) defying Ottoman power and Christian Orthodox notables. In some areas of Central Greece, their power and influence were so significant as to be recruited by the Ottomans as militiamen (*Armatoloi*), supposedly to protect the area from their alter egos, the klephts. Nonetheless, they kept reversing their allegiances according to their interests by alternating between the two roles.

During the eighteenth century, Greek merchants operating in the Balkans and the Asia Minor but also in Central and Northern Europe, including Russia, and Greek shipowners living in small islands of the Aegean and operating in the Mediterranean and the Atlantic, managed to amass great wealth since they controlled a great part of the trade in the Ottoman empire. In the eighteenth and early nineteenth century, they exploited the vacuums in French and English maritime trade during the Second Hundred Years' War (1714–1815), especially the Napoleonic Wars (1803–1815). One of the consequences of the end in European Wars in 1815 was idle capital and labor in the Greek peninsula, mostly in the islands.

The intellectual elite of the Orthodox Christians who spoke Greek was influenced by the eighteenth century enlightenment, earlier from the other elites in the Ottoman Empire. By the middle of the eighteenth century and in the first half of nineteenth century, major commercial, professional, and academic

(predominantly student) communities of Greek-speaking Orthodox Christians were formed in Vienna, Italian cities, Odessa, and other European metropolitan and commercial centers. Some clergymen and sons of wealthy merchants educated abroad were among the leading proponents of the enlightenment ideas, major works were translated into Greek, and schools teaching Greek were proliferated in key commercial centers. The backlash from the conservative ecclesiastical hierarchy after the French Revolution failed to stop the dissemination of ideas, even in mainland Greece. The half-baked Greek enlightenment undermined the authority of the Church, reconnected the Christian Orthodox elites with Ancient Greece and Western Europe, and created the fertile ground for a revolution that was not only nationalist but also democratic and liberal.

The Greek War of Independence (1821–1832)

The Greek War of Independence was not the first insurgency against the Ottoman Empire in the Balkans. Not even the first by Greek speaking people. The last major uprising by Greeks was the Orlov revolt (1770–1771), incited by Russia's Catherine the Great to distract the Ottomans during the Russo-Turkish War of 1768–1774 and it was rather easily suppressed by the Ottomans.

The French Revolution and the Napoleonic Wars were the events that triggered the Greek Revolution. Rigas Feraios, an intellectual and early revolutionary, tried to secure Napoleon's support for an insurgency in the Balkans but he was betrayed and he was arrested by the Austrians who delivered him to the Ottomans who tortured and executed him. He became the first national hero of modern Greece and he inspired similar revolutionary activities, especially in the Greek diaspora, between merchants, students, and some Phanariotes.

The Greek War of Independence was organized by unlikely revolutionaries: a group of small-time businessmen, with ties to freemasonry and Carbonarism, who founded *Filiki Etairia* (Society of Friends) as a secret organization, on September 1814 in Odessa. By 1821, Filiki Etairia had hundreds of members in the Greek peninsula, mostly in the Peloponnese, the islands and the Constantinople. The plan was to incite a revolution in the Balkans (from Bucharest to Crete) but the two major points for the break out was first the Danubian Principalities and then the Peloponnese. The leader of the revolution was Prince Alexandros Ypsilantis, a major general of the Russian army and aide-de-camp of Tsar Alexander I.

Still, leading Greek intellectuals, like the liberal Adamantios Korais and statesmen, like Count Ioannis Kapodistrias, foreign minister of Russia from 1816 to 1822, were negative to the idea. They, respectively, believed that Greeks were not ready or that the international situation was extremely hostile to revolutions after the Congress of Vienna (1814–1815).

The Ypsilantis Campaign in Moldavia and Wallachia started on February of 1821 and failed after four excruciating months. Ypsilantis was arrested and imprisoned by the Austrian authorities. He failed to obtain the support of other Balkan peoples, the uprising was renounced by Russia and the revolutionaries were excommunicated by the Patriarchate in Constantinople.

Nonetheless, Greeks took advantage of the fact that Ottomans were distracted by the uprising of Ali Pasha of Ioannina, a powerful warlord who challenged the Ottoman Empire. The suppression of his uprising occupied formidable Ottoman armies for almost a year before and a year after the breakout of the Greek revolution. These forces, thus, were unavailable to quash the insurgents in the Greek peninsula, offering to the revolutionaries the necessary breathing space. Ali Pasha's insurgency was instrumental for the decision by Filiki Etairia leadership to start the revolution during the Spring of 1821.

Even though the Greek revolution was dangerous for the *status quo* that the Congress of Vienna had established in 1815, it generated conflicting sentiments in Europe. Klemens von Metternich, the powerful Chancellor of the Austrian Empire, was very hostile and suspicious

of the Greek rebels. He thought that this was not a purely local insurgency of aggrieved Christians against the cruel and corrupted Ottomans but the sperm of a broader revolution in the Balkans, a threat to multiethnic empires, a radical uprising. He was right. The Greek War of Independence stroke a chord with liberal thinkers and activists but also with romantic poets and writers. These were the Greeks after all, fighting to liberate their "sacred" land from the "uncivilized" Muslim conquerors. The first victories of the Greeks in Peloponnese, Central Greece and the sea (several islands offered a strong makeshift fleet from converted merchant vessels), the Ottoman atrocities against Christian populations (in retaliation for the revolution but also for Greek atrocities) and the intelligent way the revolutionaries presented themselves to Europe had two impressive results: the local insurgency became, almost instantaneously, an international event and scores of influential Europeans, poets (like Lord Byron), and intellectuals (like Jeremy Bentham) decided to support the cause financially or politically. This led to an impressive Philhellenic movement in Europe which was instrumental for the final liberation of Greeks.

The protagonists of the Greek War of Independence came under broad but not necessarily mutually exclusive, categories. Two important groups were:

(a) The former klephts and armatoloi who were brave and shrewd enough to defeat the superior military forces that Ottomans sent to the south. Theodoros Kolokotronis, an ex-klepth, was the most important military leader who also tried to capitalize politically from his well-deserved fame and popularity. However, his view of the world was parochial and his behavior quite opportunistic. He failed to become the "Greek George Washington" because he lacked the American founder's humility and the political sense to understand what the stakes were.
(b) The westernized intellectuals, most of them vehicles of liberal and democratic ideas, were the ones who managed to take over the Revolution. They were former Phanariotes or children of wealthy merchants who studied in Europe and came back to join the Revolution with a clear agenda: to transform the new country into a European constitutional democracy. Their leader was Alexandros Mavrokordatos. They passed, during the Revolution, three liberal and democratic constitutions, approved by national assemblies.

These two factions came to an eventual conflict which was far more multifaceted than a clash between liberals and traditionalists. After the initial military success of the Greeks, there was a ruthless struggle for domination which led to a civil war in two stages. The Ottomans found the opportunity to regroup, asking also the help of Muhammad Ali of Egypt who sent his son Ibrahim to Greece with cavalry and infantry of 20,000 well-trained, by European officers, Egyptian troops. Ibrahim, with the help of other Ottoman forces, managed to quench the revolution in the greatest part of central and southern Greece but his cruel tactics, the massacre of Missolonghi (April 1826) and geopolitical considerations, led to the intervention of three major European powers. The warships of United Kingdom, France, and Russia destroyed Ibrahim's armada at Navarino bay in October 1827. It took another year for his forces to evacuate Peloponnese, under the pressure of a French expeditionary force. The independence of Greece was declared in February 1830 and with the London Protocol of August 1832, the Kingdom of Greece was established.

Instrumental to this result was the competition among the three Great Powers to influence the direction and the alliances of the new small state but mostly the smart policy of the British foreign minister, George Canning. He was the first to realize that Greece could be a natural and strategic ally in Eastern Mediterranean and the western-oriented Greek faction managed, in 1825, to persuade the revolutionaries to officially ask for the protection of Great Britain. From the mid-1820 (esp. after 1854) to 1947, Greece remained in the orbit of the overbearing British

Empire, being its staunchest ally but also benefitting from the strong political and diplomatic ties with the major power of the era.

Even though the intention of most Revolutionaries, as well as of the Great Powers, was for Greece to become a monarchy, Greeks elected, in 1827, as a President of the new independent state, Ioannis Kapodistrias. Kapodistrias arrived in Greece in early 1828 and realized that the only way to govern over the different factions was with a strong hand. He abolished the Constitution and he assumed dictatorial powers with the initial consent of most representatives in the revolutionary assembly. This was the end of the first Greek republic (1822–1828). However, Kapodistrias was the only Greek who could do the job (statecraft) adequately. His authoritarian reformism, especially his attempt to establish a centralized rule over local notables and warlords, led to his assassination in late September 1831. In less than 4 years, he managed to transform Greece from something resembling a state, with warlords and local notables reigning over poor farmers to a semblance of a European-oriented state. His political agenda was modernization. He organized the administration, he fought highwaymen and pirates, he organized tax authorities and the judiciary, and he tried to secure international loans. His most important legacy was land reform and the establishment of a rudimentary national education system.

Growing Pains: The New Kingdom (1833–1911)

Kapodistrias' assassination is a traumatic event in Greek history. His death led to 16 months of civil unrest and anarchy. So, when the Great Powers elected the Bavarian Prince Otto from the House of Wittelsbach for the throne of the new Kingdom in 1832, almost everyone in Greece received their decision with a relief. The 17-year-old Otto arrived in Greece in early 1833 with a Regency council who governed Greece for almost 3 years, until Otto reached majority. Both the Regency Council and Otto were autocrats. Despite the popularity of the young King, his absolutism and the fact that he was a Roman Catholic, married to the Protestant and childless Amalia, were among the causes of resentment together with austerity measures he had to adopt, including suspension of benefits to war veterans, due to Greece's insolvency. This led to a bloodless insurgency on September 1843. The uprising was supported by the military garrison of Athens and several politicians. Otto had to yield to their demands. He promised to grant a constitution and to terminate the involvement of Bavarians in the administration. The new Constitution was proclaimed in March 1844, transforming Greece into a constitutional monarchy with a bicameral parliament. With the electoral law of March 18, 1844, Greece was the first country in the world to introduce universal male suffrage – nine out of ten male adult citizens obtained voting rights. However, Otto's destabilizing constitutional transgressions and his direct involvement in politics led to more bitterness and criticism. In October 1862, he was forced to abdicate and leave Greece.

The Bavarian legacy is mixed. The three decades were a period of establishing hierarchy and bureaucracy in the public administration and the army, of organizing education and judiciary, of modernization and Europeanization. One of the regents, the law professor Georg Ludwig von Maurer, was instrumental in the process of the adoption of modern European institutions and legal codes by the new state but also of the unpopular "nationalization" of the Greek Orthodox Church, even though he stayed in office for only 18 months. Otto had moved the capital from Nafplion to Athens for obvious symbolic reasons. During his reign, he adopted the irredentist policy of the "Great Idea" (*Megali Idea*), i.e., the enlargement of the Greek state to include all lands under Ottoman rule inhabited by large Greek-speaking populations, including Constantinople, southern Balkans, the Aegean Sea islands, Crete, Cyprus, and the western part of Asia Minor. From 1844 to 1922, irredentist nationalism became the dominant policy for Greece leading to impressive territorial expansions but also to a major catastrophe. One of Otto's major failures was to take advantage of

the Crimean War (1853–1856) in order to expand against the Ottoman Empire.

Otto's abdication was the result of an uprising against his reign. His successor was a Danish Prince from the House of Glücksburg who became King George I and reigned for half a century, from 1863 to 1913. He was a solid anglophile and willing to accept a democratic constitution. With the new constitution of 1864, Greece became one of the first parliamentary democracies in the world, especially when the democratic principle was reinforced in 1875 when popular sovereignty was guaranteed by the introduction of the constitutional principle that the government should enjoy the confidence of the Parliament.

The leading reformist political leader of the late nineteenth century was the liberal Charilaos Trikoupis. From 1875 to 1895, he dominated Greek politics. He became prime minister seven times and he governed more than a decade in total. He began his career fighting the constitutional transgressions of King George I. During his tenures, the infrastructure which modernized Greece was built, but it was funded by foreign loans, leading the country to bankruptcy in 1893. A few years later, in the summer of 1896, Athens hosted the first international Olympic Games in modern history but Greece had also to face an "unfortunate" war with Turkey in 1897.

From 1864 to 1881, Greek territory was enlarged peacefully. Britain ceded the Ionian islands to Greece in 1864 as a gift and rewarded Greece with Thessaly and part of Epirus, for not siding with Russia, despite the nationalistic temptation, in the Russo-Turkish War of 1877–1878. However, these minor territorial gains couldn't satisfy Megali Idea, in an era where the Eastern Question preoccupied the minds of governments and people in the Balkans. The evolving Cretan question (Cretans revolted almost uninterruptedly against the Ottoman authorities during the nineteenth century) pressured enormously the Greek governments and kindled nationalistic feelings. Nevertheless, British protection again ensured that Greek territory remained intact, even after the Greco-Turkish War of 1897 while Crete became autonomous in 1898 under Prince George of Greece. But Greece's military defeat led to a national humiliation and the imposition of international control of Greek finances which led to an impressive fiscal consolidation.

Belle Époque was not so superb for Greece. The bitter military defeat and the bankruptcy led to the disillusionment of Greeks and resentment against the Palace. The Cretan issue was not considered resolved since Cretans themselves were not satisfied with autonomy, they wanted to be united with the "motherland." At the same time, a new antagonism begun in Ottoman Macedonia, mostly between Greeks and Bulgarians. From 1893 to 1908 (when the Young Turks Revolution took place), the two countries used armed propaganda, cultural and religious influence to attract and subdue the mixed-ethnically people living in Macedonia. Additional chronic problems, during the nineteenth century, were political corruption, rigged elections, a powerful clientelist system, and a dysfunctional bureaucracy.

The turmoil in Macedonia but mostly in Crete and the initial success of the Young Turks led to the first military coup in the twentieth century Greece. The pronunciamento was bloodless, organized by young and politically inexperienced military officers without a clear political agenda, in mid-August 1909. The government capitulated to their demands, but a political stalemate was the inevitable outcome of their indecisiveness. Finally, they offered the political leadership to Eleftherios Venizelos, a middle-aged Cretan politician and former revolutionary. Venizelos arrived in Greece and decided, reluctantly at first, to play the political game by levelling first the play field. He was a political genius and the greatest statesman in modern Greek history. He dominated Greek politics for 25 years with his impressive successes but also failures and his polarizing figure. From 1910 to 1915, he managed to gain the trust of King George I, he was appointed Prime Minister, he won election after election, he founded the first modern Greek political party (the Liberal Party), he amended the constitution (in 1911), he enforced many structural reforms in almost every area: from the reorganization and training of the army to education.

The Decade of Triumph and Tragedy (1912–1922)

The disappointment that followed the promise of the Young Turks Revolution for equality and democratic representation of the minorities and the ripening of the Eastern Question were two of the reasons behind the Balkan Wars of 1912–1913. A military alliance of Bulgaria, Serbia, Montenegro, and Greece attacked the Ottoman Empire in October 1912. Bulgarians and Serbians, with their strong armies, never believed that the relatively weak Greek army would be so successful and fast as to reach the finishing line first by capturing Thessaloniki, the bone of contention between the Balkan allies. Bulgaria was so disillusioned by the unforeseen Greek success as to attack Greece and Serbia after the first stage of the Balkan Wars. The Greek army was able to defeat the strong Bulgarian army and gain some additional territory. The greatest part (51%) of the apple of discord, Macedonia, became a part of the Greek state. This was the result of Venizelos' political and diplomatic genius. But he shared the credit with Prince Constantine, the son and heir of King George I who was also the military commander in chief. Venizelos and Constantine's relationship at the end of the Balkan wars was overcast by their disagreement as to the strategic objectives of the Greek expansion. King George managed to control his insolent son and to satisfy his successful prime minister, minimizing the cost of their divergence. He moved temporarily to Thessaloniki to symbolize the accession of the new territory to Greece. A few months before his golden jubilee in 1913, he was assassinated by a drunkard with anarchist sympathies. The box of Pandora was now open, as his son, became King Constantine I.

Greece had managed to double its territory and population by acquiring southern Macedonia, southern Epirus, most Aegean islands, and Crete. The integration of these new territories with sizable ethnic and religious minorities and the efficient administration was not an easy task for the small Kingdom. Nevertheless, for almost a century, Greeks waited impatiently for the Megali Idea to be realized and now, after all these years and many disappointments, everything signified that they were on a roll. They saw Venizelos as the one who made it happen politically, but Constantine was also extremely popular as a successful military leader and a symbol: he was the first King born in Greece and raised as a Greek-Orthodox, the one destined to restore Greece to greatness.

The confrontation between the two seems inevitable with hindsight. They both were strong-minded, but their views differed. Venizelos was a liberal, a staunch anglophile and he represented the interests of the most dynamic part of the new bourgeoisie. Constantine was a believer in the divine rights of monarchy, with populist impulses and rather unrelenting in his prejudices. But worst of all, Constantine was strongly pro-German. When the first World War I broke out, their conflict evolved into a National Schism, a traumatic experience with devastating consequences. From 1915 to 1936, Greece was bitterly divided between anti-Venizelists and Venizelists and eventually (after 1924) to royalists and republicans.

It was a conflict of foreign policy. But it was also a constitutional crisis. Venizelos insisted that Greece should enter the Great War on the side of Entente but Constantine was adamantly against, he favored neutrality but he also schemed in the back rooms for the benefit of the Central Powers. Venizelos resigned but even though he triumphed at the national elections, Constantine didn't relent. He saw foreign policy as his royal prerogative. Venizelos resigned for the second time. Entente pressured ruthlessly the royal governments that followed to submit to its demands and, when German-escorted Bulgarian troops seized part of the Greek Macedonia, Venizelos decided to set up a provisional government in Thessaloniki with the support of the French army. Greece was torn. There was even an armed confrontation in the streets of Athens between the army of the royalist government and French forces in November 1916, which led to an even deeper wedge between royalists and Venizelists. A naval blockade by the Allies made Constantine leave Athens (he didn't abdicate) together with his first-born son and heir, George, in June 1917. Venizelos returned to Athens and assumed power. Alexander, the second-born son of Constantine, became a kind of "interim" King. He was ideal for the job. A malleable individual, he was easily handled by

Venizelos. Greece entered the war on the side of the Allies at the last stage of it helping them in their offensive in May 1918 that led to the capitulation of Bulgaria.

The fact that Bulgaria but also the Ottoman Empire sided with the Central Powers who lost the war gave Greece an enormous diplomatic advantage. Venizelos reinforced this advantage by sending Greek troops, in 1919, to fight the Red Army together with the White Forces at the multinational military expedition in Ukraine. Venizelos proved himself the most credible ally in the Balkans and he was splendidly rewarded in the Treaty of Sèvres of 1920. Western and Eastern Thrace were delivered to Greece. The Greek army reached the outskirts of Constantinople and landed in western Asia Minor, Smyrna (today Izmir) and a large surrounding area with the objective to annex it to Greece, with the approval of the Allies, after a referendum.

Venizelos could not enjoy his well-deserved triumph. King Alexander died unexpectedly and himself was nearly killed in an assassination attempt by two royalist officers. The elections his government held in November of 1920 to capitalize politically on his successes led to a defeat. A disheartened Venizelos left Greece for France. The exultant royalists organized a plebiscite to bring back Constantine, despite stern warnings by the Allies.

An ailing and less confident Constantine and his incompetent governments ruled Greece for the next 2 years taking revenge against the Venizelists and trying to cajole the British. Their foolish blunder was a futile attempt to destroy the Turkish forces led by Mustafa Kemal (later Atatürk) by reaching and capturing Ankara. This was a strategic mistake comparable to the French and German invasion of Russia in the nineteenth and twentieth century. After 2 years of pyrrhic victories, the Greek army was defeated and retreated, leaving Smyrna to the advancing Turkish army. Smyrna was destroyed, and its Greek inhabitants were killed, captured, fled, or deported. The, more than 2,500 years, Greek presence in Asia Minor ended with the greatest catastrophe in Modern Greek history. It was also the end of Megali Idea. Greece had to abandon eastern Thrace, Smyrna, and two small Aegean islands. Greeks left their ancestral homelands in Pontus and Cappadocia.

Bitter Divide: The Interwar Period (1922–1940)

After the catastrophe, a military coup, organized by mostly Venizelist officers, made Constantine abdicate in favor of his firstborn son who became King George II, but only temporarily. The process of transforming Greece into a republic had started. The radical Venizelist officers organized a mock martial court trial of the leadership of the royalists. Six of them were sentenced to death and they were speedily executed. Their execution led to political scars that for several decades plagued Greece. The royalists abstained in the 1923 national elections and the result was the political dominance of republican liberals for the next decade. On March 25, 1924, the second Greek Republic was proclaimed and a new constitution was adopted in 1927, but most royalists didn't accept the new regime and the National Schism deepened. To make matters worse, one radical republican, General Theodoros Pangalos, assumed dictatorial powers for a year. One of the challenges was to accommodate 1.3 million refugees from Asia Minor who integrated themselves completely only after decades.

Venizelos returned to power after a landslide electoral victory in 1928. His tenure from 1928 to 1932 was one of the most fruitful in almost all areas. From structural and institutional reforms to international relations. Even though the refugees were a great part of his electorate, he dared to visit Ankara and sign a friendship agreement with Atatürk, ending a century of conflict. However, the Great Depression led Greece to economic destabilization and to the loss of majority support to Venizelists. The anti-Venizelists ("People's Party") came back to power with a vengeance in the 1933 elections, ending the dominance of the Liberal Party and eventually capturing the bureaucracy and the military. Venizelos (dismayed after a second attempt against his life) and the republican officers tried to prevent the restoration of the King with an ill-executed coup in 1935 which led Venizelos to an

exile in Paris where he died the following year. George II returned with a rigged plebiscite. This was the end of the second Greek Republic (1924–1935). The Liberals accepted the regime change, but it was not enough. The rise of the Greek Communist Party and a wave of labor strikes was used as pretext by George II to accept a dictatorship under Ioannis Metaxas, an experienced and competent but ruthless retired officer and royalist politician. The Greek people tired by almost three decades of wars and political turmoil didn't resist to the new regime which was developed into a conservative authoritarian government fashioned after Fascist Italy.

The Decade of Wars (1940–1949)

Despite the rise of Nazi Germany, George II and Metaxas had learned their lesson from World War I. They had decided to keep Greece neutral, if possible, but if there was no choice, to remain faithful to Britain. When Mussolini, in October 1940, invaded Greece, the King and Metaxas decided to fight back, the Greek people totally agreeing with their decision and fighting bravely. The Greek army not only fended Italian forces off but also it managed to humiliate Mussolini by advancing in the under Italian control Albania capturing one city after another. This was the first military success against the Axis, so Hitler had to intervene while he prepared Operation Barbarossa against the Soviet Union. Germans invaded Greece in April 1941. The exhausted Greek army and a small British expeditionary force was not able to protect the country. Nevertheless, it took Germans almost 2 months to occupy the whole of Greece due to the fierce resistance by Greek and British forces in Crete.

King George (Metaxas died unexpectedly in early 1941) fled to Egypt where he appointed a new prime minister, the former Venizelist, Ioannis Tsouderos. In Athens, a puppet regime was established by collaborationists. Greece was separated in three different zones, occupied respectively by German, Italian and Bulgarian forces. Several resistance groups were established from the very beginning in mountainous areas with the help of the British. The most well organized was EAM, that was controlled by the Communist Party with the same agenda as similar partisan organizations in occupied Eastern Europe, i.e., to distract German forces in their invasion of Soviet Union and to seize power after the end of the war.

EAM and its military branch, ELAS, managed to annihilate almost every rival resistance group and dominate the field. When the defeat of Germany was more than obvious, the puppet regime tried to organize paramilitary groups to prevent the capture of power by the communists after the German retreat. EAM formed a provisional government in order to impose its participation to the government in exile. A compromise was achieved with the backing of the British and a coalition government was formed under a seasoned liberal politician and a former associate of Venizelos, George Papandreou. Papandreou returned to Athens after the departure of German forces in October 1944 but EAM/ELAS controlled the rest of the country. The communists could have grabbed power rather easily. However, Winston Churchill had already secured Greece for the West in his negotiation with Stalin and the latter didn't encourage Greek communists who were baffled. The behavior of EAM ministers in the coalition government was not cooperative and the noncommunist members didn't trust them. EAM didn't want to disarm its army and relinquish the control of the periphery.

After the bitter resignation of its ministers, EAM organized a strong demonstration in early December 1944. The demonstration had a wretched ending after police killed several demonstrators. This led to a bloody confrontation in the streets of Athens: The numerically superior EAM/ELAS tried to seize control of the capital against the British forces stationed in Athens which were supported by a coalition faithful to the government: veterans from the Greco-Italian war and the North Africa Campaign, right-wing guerilla groups, policemen, even some former collaborationists. The communist forces were far superior, but they couldn't defeat this unlikely pro-government coalition. Winston Churchill ordered troops from the Italian front to join the small contingent in Athens and quell the

rebellion. Under considerable domestic pressure, the British prime minister deemed the situation critical enough for him to spend Christmas day in Athens, in an inconclusive mediation effort. Overpowered, the communists capitulated and accepted the disarmament of ELAS.

However, atrocities from both sides (Red/White terror), the decision of the Communist Party to abstain from the first postwar national elections in March 1946 (leading to a triumph for the royalist Right) and the return of King George II after a referendum polarized even more the Greeks and led to the final stage of the Civil War (1946–1949). This time the Communist Party had decided to fight to the end. The prospects were good: the communists controlled many parts of Greece, the kindred governments of Albania, Yugoslavia, and Bulgaria already provided material support and safe haven for the communist army, Soviet aid was anticipated while the exhausted British Empire relinquished its traditional role in Greece. But, it was replaced immediately by the United States who decided to assert its role as the postwar superpower, providing abundant economic and military support to the Greek government, under a veteran leader of the minority Liberals, Themistoklis Sofoulis. The failure of the communists to capture and hold any significant town was exacerbated by the Tito-Stalin split which eventually led to Yugoslavia closing its borders. After two inconclusive years of fighting, in summer 1949, the government forces launched their final assault against the communist stronghold on the mountains of northwestern Greece. By the end of August, the remnants of the defeated communist forces fled to Albania.

From Illiberal Democracy to Dictatorship (1949–1974)

Greece remained part of the western democratic world (joining NATO in 1952), but the price was an illiberal (tutelary) democracy which treated the defeated communists (and fellow travelers in general) harshly, deepening the rift instead of investing in reconciliation. The United States exerted an enormous influence on Greek politics while the King (Paul, the third son of Constantine I succeeded George II in 1947 and Paul's son Constantine II succeeded his father in 1964) antagonized elected governments, especially during the mid-1960s. Nevertheless, Greece remained a democracy. The Communist Party was banned but it was represented by a leftist party, a political front, which, 9 years after the end of the Civil War, managed to become, for a brief period, the second strongest party in parliament. The leading politician of the period was Konstantinos Karamanlis, a conservative reformist who in 8 years (1955–1963) transformed Greece. He successfully pursued Greece's association with the European Communities and he was responsible for the most spectacular economic growth due to rapid industrialization and investment in infrastructure and tourism. Karamanlis also managed to reach a compromise solution for the Cyprus problem which since the early 1950s had plagued Greece's foreign relations and domestic politics. Greek Cypriots were the majority (78%) in the, under British administration, island. Nonetheless, there was a sizable Turkish Cypriot minority (18%) who felt threatened by the prospect of *Enosis* (the union of the island with Greece). Rather than insisting on Enosis, Greece agreed with Turkey and Great Britain for Cyprus to become an independent Republic in 1960. After a quarrel with King Paul, Karamanlis resigned and then lost the elections of 1963 to a centrist party led by George Papandreou. Papandreou himself had to resign in the summer of 1965 when the young King Constantine denied him his constitutional prerogative to discharge the defense minister of his government.

A period of political turmoil led to a military coup, organized by colonels against the political system in general in April 1967. Constantine II fled the country 8 months later, after a failed counter-coup, and all political activity ceased for 7 years. The self-confident leader of the military junta, George Papadopoulos, proclaimed Greece a "republic" in 1973, appointed a docile civilian government, and announced (controlled) elections for the following year. His plans went awry after student riots in November 1973 gave to the

G

regime hardliners an excuse to sack Papadopoulos and forestall the supposed liberalization process. Their foolish mistake was the attempt to assassinate the President of the Republic of Cyprus, Makarios, in order to annex Cyprus to Greece in mid-July of 1974. Turkey invaded Cyprus (the north of the island is still occupied by Turkish forces) and the chain of events led to the fall of the humiliated military junta. Konstantinos Karamanlis returned on July 24, 1974, to Greece after a self-imposed exile of 11 years.

The Third Greek Republic: From Metapolitefsi to the Economic Crisis (1974–)

Karamanlis, as the prime minister in a coalition government, managed to restore democracy (*metapolitefsi*) in a paradigmatic way, he legalized the Communist Party, and he passed several necessary structural reforms. Greece became a Republic with the constitution of 1975 (currently in force) after the 1974 referendum which ended Greek monarchy. It was the beginning of the current period of the Third Greek Republic which is the less tumultuous in modern Greek history. Karamanlis governed from 1974 to 1980 when he was elected President of the Republic. In January 1981, Greece became a full member of the European Communities.

The 1980s were dominated by the socialist party PASOK under its charismatic leader Andreas Papandreou (son of George) whose welfare populism triumphed. He governed Greece from 1981 to 1996 with a short interval (1989–1993) when the conservative party of Karamanlis (New Democracy) returned to power with a liberal reformist leader, Constantine Mitsotakis. PASOK's dominance continued after the death of Papandreou with the reformist social democrat, Kostas Simitis, as prime minister. Simitis' 8-years tenure was marked by high growth rates and Greece's accession to the Eurozone but with no major structural reforms. He was succeeded by Kostas Karamanlis (the nephew of the former Prime Minister and President) whose free-spending policies precipitated the Greek debt crisis of 2009–2010.

Karamanlis lost the 2009 elections to PASOK's George Papandreou (son of Andreas) who had to deal with the sovereign debt crisis. In May 2010, Greece entered a bailout agreement with the Eurozone countries and the IMF. Two more bailouts followed since, reaching loans to Greece to a total of over €250 billion. The adopted austerity measures agreed between Greece and its creditors (three economic adjustment programs) led to a remarkable fiscal consolidation but they stagnated the economy. This was so, because mostly fiscal measures were adopted and only some half-baked institutional reforms were enforced. The first bailout agreements were enforced by PASOK and then two coalition governments formed by New Democracy and PASOK. At the end of 2014, Greece was in a process of fragile recuperation which was upset by the victory of the radical leftist and populist SYRIZA. SYRIZA tried to renegotiate the bailout agreement by following a self-defeating brinkmanship strategy which led to its bitter capitulation in July 2015. The third bailout agreement included austerity measures, harsher than the preceding ones. Since then SYRIZA and its leader Alexis Tsipras managed to secure their power by cooperating with the creditors in a government coalition with a minor ultra-right-wing party. Despite its obvious failure to deliver on its electoral promises to get rid of austerity, this coalition outlived all preceding bailout governments.

Greece in the Twenty-First Century

Modern Greece has a history of almost two centuries. During these centuries, the country managed to move from the backwaters of Europe to a prosperous liberal democracy. From 1929 to 1980, Greece had an average annual rate of growth of income per capita of 5.2% (during the same period, Japan had an average of 4.9% and Germany 3.0%). However, this development was based on extractive institutions. The membership in the European Union and the Eurozone helped Greece put its extractive institutions under the rug of EU convergence funds, cheap international borrowing and fudged statistics. The crisis of

2008 was the triggering effect of the perfect storm that hit Greece in early 2010. Greece must replace its extractive institutions with inclusive institutions suitable for economic growth. This should be the new "Megali Idea" for the Greek people.

Acknowledgments The author is grateful to Konstantina Botsiou, Maria Efthymiou, Yulie Foka-Kavalieraki, Basil Gounaris, Evanthis Hatzivassiliou, Akritas Kaidatzis, Dimitris Keridis, Nikos Marantzidis, Iakovos Michailidis, Ioanna Sapfo Pepelasis, Ioannis Stefanidis, Thanos Veremis, Spyros Vlachopoulos and Elpida Vogli for valuable comments to a previous draft. The usual caveat applies.

References

Brewer D (2001) The flame of freedom: the Greek War of Independence, 1821–1833. John Murray, London
Brewer D (2010) Greece, the hidden centuries: Turkish rule from the fall of Constantinople to Greek Independence. I.B. Tauris, London
Brewer D (2016) Greece, the decade of war: occupation, resistance and civil war. I.B. Tauris, London
Clogg R (2013) A concise history of Greece, 3rd edn. Cambridge University Press, Cambridge
Diamandouros N (1972) Political modernization, social conflict and cultural cleavage in the formation of the modern Greek state: 1821–1828. Columbia University PhD thesis
Featherstone K (ed) (2014) Europe in modern Greek history. Hurst, London
Featherstone K, Papadimitriou D (2015) Prime ministers in Greece: the paradox of power. Oxford University Press, New York
Gallant TW (2015) The Edinburgh history of the Greeks, 1768 to 1913 (the long nineteenth century). Edinburgh University Press, Edinburgh
Gallant TW (2016) Modern Greece: from the war of independence to the present, 2nd edn. Bloomsbury, London
Gounaris BG (1996) Social cleavages and national 'awakening' in Ottoman Macedonia. East Eur Q 29:409–426
Greene M (2015) The Edinburgh history of the Greeks, 1453 to 1768 (the Ottoman Empire). Edinburgh University Press, Edinburgh
Hatzis AN (2002) The short-lived influence of the Napoleonic Civil Code in 19th century Greece. Eur J Law Econ 14:253–263
Hatzis AN (2018) Greece's institutional trap. Manag Decis Econ 39. (forthcoming)
Hatzivassiliou E (2011) Greece and the Cold War: front line state, 1952–1967. Routedge, London
Hering G (1992) Die politischen Parteien in Griechenland 1821–1936, (2 vols.). R. Oldenbourg Verlag, Munich
Iatrides JO (1972) Revolt in Athens: the Greek Communist Second Round, 1944–1945. Princeton University Press, Princeton
Kaltchas N (1940) Introduction to the constitutional history of modern Greece. Columbia University Press, New York
Kalyvas SN (2015) Modern Greece: what everyone needs to know. Oxford, New York
Keridis D (2009) Historical dictionary of modern Greece. Scarecrow Press, Lanham
Kitromilides PM (ed) (2006) Eleftherios Venizelos: the trials of statesmanship. Edinburgh University Press, Edinburgh
Kitromilides PM (2013) Enlightenment and revolution: the making of modern Greece. Oxford University Press, Oxford
Koliopoulos JS, Veremis TM (2007) Greece: the modern sequel (from 1821 to the present), 2nd edn. Hurst, London
Koliopoulos JS, Veremis TM (2009) Modern Greece: a history since 1821. Wiley-Blackwell, Malden
Kostis K (2018) History's spoiled children: the story of modern Greece. Oxford University Press, Oxford
Mavrogordatos GT (1983) Stillborn republic: social coalitions and party strategies in Greece, 1922–1936. University of California Press, Berkeley
Mazower M (1993) Inside Hitler's Greece: the experience of occupation, 1941–44. Yale University Press, New Haven
Mazower M (ed) (2000) After the war was over: reconstructing the family, nation, and state in Greece, 1943–1960. Princeton University Press, Princeton
Mazower M (2004) Salonica, city of ghosts: Christians, Muslims and Jews 1430–1950. Vintage Books, New York
McNeil WH (1978) The metamorphosis of Greece since World War II. University of Chicago Press, Chicago
Pappas TS (1998) Making Party Democracy in Greece. Palgrave
Petropoulos AJ (1968) Politics and statecraft in the kingdom of Greece: 1833–1843. Princeton University Press, Princeton
Ricks D, Beaton R (eds) (2016) The making of modern Greece: nationalism, romanticism, and the uses of the past (1797–1896). Routledge, London
Stefanidis ID (2007) Stirring the Greek nation: political culture, irredentism and anti-Americanism in post-war Greece, 1945–1967. Ashgate, Aldershot
Woodhouse CM (1976) The struggle for Greece, 1941–1949. Ivan R. Dee, Chicago
Woodhouse CM (2000) Modern Greece: a short history. Faber & Faber, London

G

Greenhouse Effect

► Global Warming