© The Author(s) 2022. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Empirical research in comparative constitutional law: The cool kid on the block or all smoke and mirrors?

Niels Petersen* and Konstantin Chatziathanasiou**

In recent years, we have observed an explosion of empirical research in the field of comparative constitutional law. This contribution seeks to evaluate the current state of affairs. We analyze select studies from four different areas: the consequences of and the reasons for constitutional design choices, as well as the diffusion and the effectiveness of constitutional rights. We find that the empirical identification strategies of many of the analyzed studies have significant weaknesses, and inspire only limited confidence in their results. Nevertheless, we argue that empirical research in comparative constitutional law is of fundamental importance. It not only draws our attention to issues that would otherwise elude our view, but also gives us a chance to refine the methodology in order to develop better strategies to answer some of the decisive questions preoccupying comparative constitutional law scholarship.

1. Introduction

In his influential book, *Comparative Matters*, Ran Hirschl argued that comparative constitutional law should rely to a much greater extent on social science insights and social science methodology.[1] His call was heard, it seems, as in recent years we have

* Co-director, Institute for International and Comparative Public Law and Professor of Public Law, University of Münster, Münster, Germany; Research Affiliate, Max Planck Institute for Research on Collective Goods, Bonn, Germany. Email: niels.petersen@uni-muenster.de.

** Postdoctoral Research Fellow, Institute for International and Comparative Public Law, University of Münster, Münster, Germany; Visiting Researcher, Max Planck Institute for Research on Collective Goods, Bonn, Germany. Email: kchatzia@uni-muenster.de.

The authors thank Stefanie Egidy, Jens Frankenreiter, and Samuel Issacharoff for valuable comments on earlier drafts of this article.

[1] Ran Hirschl, Comparative Matters: The Renaissance of Comparative Constitutional Law 151–91 (2014). For pointed critiques of Hirschl's proposal as a predominant approach to comparative constitutional law, see Armin von Bogdandy, *Comparative Constitutional Law as a Social Science? A Hegelian Reaction to Ran Hirschl's Comparative Matters*, 49 Verfassung und Recht in Übersee 278 (2016); Theunis Roux, *Comparative Constitutional Studies: Two Fields or One?*, 13 Ann. Rev. L. Soc. Sci. 123 (2017).

I•CON (2021), Vol. 19 No. 5, 1810–1834 https://doi.org/10.1093/icon/moab125

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

been able to observe a proliferation of empirical studies in comparative constitutional law. These studies rely mostly on quantitative techniques of analysis to investigate such diverse matters as the effect of constitutional design choices on the endurance of constitutions and the diffusion and effectiveness of constitutional rights. In this article, we want to take stock of this literature and have a closer look at its methodological approaches. We argue that many of these studies do not allow for robust conclusions regarding causal mechanisms. Currently, their value lies rather in sparking important debates on the assumptions on which many of our discussions in comparative constitutional law rest.

Our article consists of four sections. Section 2 provides a basic introduction to quantitative empirical legal research. It identifies certain methodological challenges and discusses how to address them. Section 3 analyzes the literature on the institutional design choices that are made in constitutions. There is research both on the consequences of design choices and on the reasons why specific design features were included in the constitution. Section 4 looks at the empirical literature regarding constitutional rights, which mainly deals with the diffusion and effectiveness of individual rights enshrined in constitutions. Neither section aims at providing a comprehensive review of the empirical research on comparative constitutional law. Instead, they concentrate on select studies that we consider to be particularly important and influential. Section 5, finally, concludes by analyzing how the challenges that are particular to empirical research in constitutional law could be addressed.

2. Empirical research design

Empirical legal research is not a new phenomenon. In the United States, "Empirical Legal Studies" has been an active and growing field of research for more than two decades. In this section, we take a closer look at the underlying methodological approach. In a first step, we introduce the basic intuition behind it (Section 2.1), and then, in a second step, we identify the main challenges and discuss how they can be addressed (Section 2.2).

2.1. Empirical approach to constitutional law research

Empirical legal research aims at describing and explaining legal phenomena by using data.[2] Imagine we study the relationship between the design of the political system and democratic stability, and we find that parliamentarian systems are more resilient to democratic erosion than presidential systems.[3] This correlation between parliamentarianism and democratic stability is a descriptive finding. However, it does not tell us anything about the reasons *why* this correlation exists. Descriptive research in this

[2] More generally on the distinction (and divide) between interpretative and explanatory social science research, see Alexander Rosenberg, Philosophy of Social Science (2015).

[3] The evidence is discussed in Tom Ginsburg & Aziz Huq, How to Save a Constitutional Democracy 176–87 (2018).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

sense tries to find patterns without seeking to explain them. By contrast, explanatory social science research aims at establishing causal relationships. It seeks to explain why specific patterns can be observed and which factors caused them. Does parliamentarianism have a positive, causal effect on democratic stability in our hypothetical example? Or is the latter caused by other factors?

The concept of causality employed in the social sciences differs markedly from the understanding of causality prevalent in legal thought. Lawyers typically have a deterministic understanding of causality. Criminal lawyers are concerned with the question whether a suspect caused the death of a victim. Tort lawyers are interested in whether a particular action caused a specific damage. By contrast, social scientists have a probabilistic understanding of causality. They are interested in general regularities and seek to understand whether the presence of a factor *X* makes the occurrence of a factor *Y more likely*. Let us assume that we are interested in the relationship between democracy and economic performance.[4] Applying a deterministic understanding of causality, we would inquire whether democracy always has a positive effect on economic performance. With a probabilistic understanding, we would inquire whether democracy makes a higher level of economic growth more likely. This implies that, even if democracy had a positive effect on economic performance, it would nonetheless be possible that some autocracies have a higher level of economic growth than many democracies.[5]

Methodologically, there are quantitative and qualitative approaches to empirical research. The difference lies in the representation and processing of the underlying data. Qualitative studies rely on non-numeric, typically verbal, information; while quantitative studies recur to measurable and quantifiable indicators. This quantification is operationalized by consolidating complex information into a number or a string of numbers.[6] For example, the Polity IV index on the quality of democracy assigns, for each year and each state contained in the database, a number between +10 (full democracy) and –10 (full autocracy) that is supposed to designate the quality of democracy.[7]

4 This question has been subject to numerous empirical studies; see, e.g., Kenneth A. Bollen, *Political Democracy and the Timing of Development*, 44 Am. Soc. Rev. 572 (1979); Kenneth A. Bollen & Robert W. Jackman, *Economic and Noneconomic Determinants of Political Democracy in the 1960s*, 1 Res. in Pol. Soc. 27 (1985); Larry Diamond, *Economic Development and Democracy Reconsidered*, *in* Reexamining Democracy: Essays in Honor of Seymour Martin Lipset 93 (Gary Marks & Larry Diamond eds., 1992); John Benedict Londregan & Keith T. Poole, *Does High Income Promote Democracy?*, 49 World Pol. 1 (1996); Robert J. Barro, *Democracy and Growth*, 1 J. Econ. Growth 1 (1996); Robert J. Barro, *Determinants of Democracy*, 107 J. Pol. Econ. 158 (1999); Adam Przeworski, Michael E. Alvarez, José Antonio Cheibub & Fernando Limongi, Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990 (2000); Carles Boix & Susan C. Stokes, *Endogenous Democratization*, 55 World Pol. 517 (2003); Torsten Persson & Guido Tabellini, *Democracy and Development: The Devil in the Details*, 96 Am. Econ. Rev. 319 (2006); David L. Epstein et al., *Democratic Transitions*, 50 Am. J. Pol. Sci. 551 (2006); Daron Acemoglu & James A. Robinson, Economic Origins of Dictatorship and Democracy (2006); Daron Acemoglu, Simon Johnson, James A. Robinson, & Pierre Yared, *Income and Democracy*, 98 Am. Econ. Rev. 808 (2008).

5 *See* Przeworski, Alvarez, Cheibub, & Limongi, *supra* note 4, at 176–8.

6 Lee Epstein & Gary King, *The Rules of Inference*, 69 U. Chi. L. Rev. 1, 81 (2002).

7 Monty G. Marshall & Keith Jaggers, Polity IV Project: Political Regime Characteristics and Transitions, 1800–2007 (2009).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

This quantification comes at the cost of reducing information, but brings the benefit of allowing for the processing of the data through statistical techniques of analysis. By contrast, qualitative studies contain a greater wealth of information per data point, as information does not have to be reduced to one single number. Qualitative data may, for example, be collected through interviews or historical case studies. The other side of the coin is that qualitative studies usually have to rely on less available data, as our cognitive ability to collect and process such data is limited.[8]

As this discussion of strengths and weaknesses already suggests, there is no hierarchy between quantitative and qualitative studies, in the sense that quantitative studies were superior to qualitative ones.[9] The preferable approach depends rather on the research question. If the main concepts of analysis are easily quantifiable, a quantitative approach is usually preferable because it allows us to process more data. However, the less accessible the concepts of interest are to quantification, the greater the advantage of qualitative studies. Furthermore, the approaches are not mutually exclusive. For some questions, it may even be recommendable to combine qualitative and quantitative approaches. Despite these qualifications, this contribution will mainly focus on quantitative studies. The reason for this restriction is a pragmatic one: the most influential empirical studies in the field of comparative constitutional law employ a quantitative framework, so that we will also predominantly focus on quantitative techniques of analysis.

2.2. Challenges to quantitative empirical research

a) The omitted-variable bias

The ideal research design for causal-explanatory science is the controlled experiment.[10] In an experiment, participants are randomly assigned to two or more groups: the control group and the experimental group(s). Ideally, the only difference between these groups is that the experimental group receives a specific treatment that the control group does not receive. If we observe a statistically significant difference between these two groups, we can assume that the difference was caused by the treatment. Let us take a clinical trial as an example: Patients are randomly assigned to two groups. One group receives a medication against high blood pressure, while the other only receives a placebo. If, after the treatment, the blood pressure of the patients receiving the medication has decreased further than the average values of the patients receiving the placebo, and if this difference is statistically significant, we can assume that the medication has had a positive causal effect on the blood pressure of the patients

[8] Lisa Webley, *Qualitative Approaches to Empirical Legal Research*, *in* The Oxford Handbook of Empirical Legal Research 927, 934 (Peter Cane & Herbert M. Kritzer eds., 2010).

[9] *See similarly*, David S. Law, *Constitutions*, *in* The Oxford Handbook of Empirical Legal Research, *supra* note 8, at 376, 389.

[10] On the role of experiments in the social sciences and particularly in economics, see Klaus M. Schmidt, *The Role of Experiments for the Development of Economic Theories*, 10 Perspektiven der Wirtschaftspolitik 14 (2009).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

because receiving the medication was the only systematic difference between the two groups.

There may, of course, be other differences between individual members of the two groups. Some might have a better physical precondition than others, which might positively affect their ability to recover. However, if the two compared groups are sufficiently large, such differences should even out, so that they should not play a significant role in the comparison. Nevertheless, in order to avoid selection effects and other confounds, it is important that the control group and the experimental group were assigned randomly. If, for example, the medicine is given to the patients who came to an early appointment and the placebo to the patients who preferred a later appointment, we cannot exclude that the preferences for an early or late appointment are correlated with some characteristic that might influence recovery. Therefore, experiments in which the condition of random assignment is violated may only show spurious correlations and thus fail to demonstrate causality.

However, experimentation is not a feasible research method for many questions that typically preoccupy constitutional law scholars.[11] If we want to analyze whether presidential or parliamentary democracies are more stable, it is not possible to assign these conditions randomly to two groups of countries. Rather, institutional arrangements are preexisting. Therefore, if we find that presidential or parliamentarian systems are more stable, then these observations might not be caused by the institutional arrangement, but rather by factors that made countries adopt these institutional arrangements in the first place.

If experimentation is not possible, social scientists usually recur to some form of a multivariate regression analysis in order to estimate the relationship between variables. A regression analysis in itself does not allow for causal claims, so that social scientists try to control for selection effects and other confounding factors through a statistical analysis *ex post facto*. The main challenge is to avoid an omitted-variable bias.[12] If we observe a correlation between two variables, we cannot automatically conclude that one variable causally influenced the other.[13] Instead, the correlation may also be due to both variables being causally influenced by a third variable, for which the analysis did not account.

The challenge of the omitted-variable bias plays a significant role in quantitative comparative constitutional law. Let us revisit our hypothetical example of

[11] On the existing potential of experiments, however, in particular for testing behavioral assumptions underlying constitutional theory, see Stefan Voigt, *Positive Constitutional Economics: A Survey*, 90 Pub. Choice 11, 20–2 (1997); Stefan Voigt, *Empirical Constitutional Economics: Onward and Upward?*, 80 J. Econ. Behav. & Org. 319, 327–8 (2011); Konstantin Chatziathanasiou, Verfassungsstabilität: Eine Von Artikel 146 Grundgesetz Ausgehende Juristische Und (Experimental-)ökonomische Untersuchung 141–50 (2019); on the potential of experiments for legal scholarship more generally, see Colin Camerer & Eric Talley, *Experimental Study of Law and Economics*, *in* 2 Handbook of Law and Economics 1619 (A. Mitchell Polinsky & Steven Shavell eds., 2007); Christoph Engel, Legal Experiments: Mission Impossible (2013).

[12] On the challenge of the omitted-variable bias in the context of quantitative comparative law, see Mathias M. Siems, *Statistische Rechtsvergleichung*, 72 Rabels Zeitschrift für ausländisches und internationales Privatrecht 354, 372–5 (2008).

[13] James A. Davis, The Logic of Causal Order 24–7 (1985).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

comparing the stability of presidential and parliamentarian political systems. If we find that presidential systems are less stable, this instability might be caused by presidentialism. However, it might also well be that there is a third factor, for example the dominance of a specific elite, which led to the adoption of a presidential system, but at the same time also causes the instability—independently of the constitutional design choice. The correlation between presidentialism and instability would then be spurious.

b) Techniques to address the omitted-variable bias

Quantitative social sciences have developed techniques to address the omitted-variable bias. The simplest technique is the inclusion of all potentially confounding factors as control variables into the regression analysis. The problem of this approach is that we rarely know whether we have indeed included all relevant confounding factors.[14] But even if we were aware of all possible confounding factors, some of them might be difficult to observe and/or to quantify. One factor that may be relevant in many studies relating to comparative constitutional law is the prevailing culture of a state. But how do we quantify culture? In the social sciences, it is common to use the percentage of the three supposedly most widespread religions (Catholicism, Protestantism, Islam) as a proxy for culture.[15] But this approach provokes several questions: Can culture really be exclusively equated with religion? Even if that is the case, is the percentage of religious groups a good proxy? What about states where the elites stem from a religious minority? Finally, it seems to be a questionable oversimplification to conflate such different religions as Judaism, Buddhism, the Bahá'í faith or even the absence of religious beliefs into one single category, namely, other religions.

Because of the problems in addressing the omitted-variable bias simply by including control variables into the regression analysis, social scientists recur to alternative techniques. One alternative is the so-called propensity score matching technique.[16] The intuition behind matching is the following: We divide the sample of countries into two groups by finding close matches. If two countries are exactly identical, except with regard to the treatment that we would like to analyze, then the difference in the outcome has to be due to the treatment. Let us return to our example of governmental systems. If we find two countries that are exactly identical, except that one country has a presidential system while the other has a parliamentarian one, we can assume that any difference regarding the stability of the political system is not due to external factors (because these are identical), but due to either the presidential or the parliamentarian form of government.

[14] Daron Acemoglu, *Constitutions, Politics and Economics: A Review Essay on Persson and Tabellini's the Economic Effects of Constitutions*, 43 J. Econ. Literature 1025, 1029 (2005).

[15] *Seminally* Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shleifer, & Robert Vishny, *The Quality of Government*, 15 J. L., Econ. & Org. 222, 233 (1999).

[16] On propensity score matching, see Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 Biometrika 41 (1983); Joshua Angrist & Jörn-Steffen Pischke, Mostly Harmless Econometrics: An Empiricist's Companion 69–91 (2009).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

However, while this technique might look compelling in theory, it encounters similar problems as the simple inclusion of control variables.[17] Two states will never be exactly identical. Instead, we have to identify relevant factors, based on which to compare the states with regard to their level of shared characteristics. But when identifying these relevant factors, we face the same challenge that we encounter when searching for adequate control variables: Have we indeed identified *all* relevant factors? What do we do about the factors that might be relevant but are impossible or difficult to observe, such as culture?

A further technique addressing the omitted-variable bias is the instrumental variables (IV) estimation.[18] The intuition behind this technique is the following: An instrumental variable needs to be correlated with one of the explanatory variables, but must not have an independent effect on the dependent variable nor be correlated with the error term, namely, an omitted variable. In such a situation, we can estimate the effect of the independent variable on the dependent variable by measuring the effect of the instrumental variable on the dependent variable. Because the former has no independent effect on the latter, any effect has to be mediated via the independent variable.

Let us again consider an example. Imagine we want to measure the effect of the economic performance of a country on its level of democracy. Are rich countries more likely to be democratic than poor ones? We can see that this analysis is fraught with problems. On the one hand, there are many potentially confounding factors that may influence both the economic performance and the level of democracy. On the other hand, if we find a correlation between the two variables, we do not know in which direction causality runs. Does democracy cause economic growth, or is it the other way around? One potential instrumental variable might be whether a country is landlocked. We can easily see how access to a coast is important for the economic development of a state. However, it seems implausible that the variable has a direct, independent effect on the level of democracy. Therefore, any effect of the status of being a landlocked state has to be mediated through economic performance, which would allow us to measure the effect of economic performance on the level of democracy. While IV estimation is a sound statistical technique that effectively addresses the omitted-variable bias, it comes with a major drawback: It is extremely difficult to find valid instruments and to exclude that specific instruments are indeed uncorrelated with the error term and do not have a direct effect on the dependent variable.

While the three discussed techniques of dealing with the omitted-variable bias are the most common in cross-country panel data analyses, the list is not exhaustive. Other possible techniques are quasi-experiments,[19] simulations,[20] or a combination

[17] *See* Acemoglu, *supra* note 14, at 1029; Holger Spamann, *Empirical Comparative Law*, 11 Ann. Rev. L. Soc. Sci. 131, 140–1 (2015); for a nuanced critique of matching techniques based on a comparison with experimental evidence, see Kevin Arceneaux, Alan Gerber, & Donald P. Green, *Comparing Experimental and Matching Methods Using a Large-Scale Voter Mobilization Experiment*, 14 Pol. Analysis 37 (2006).

[18] *See* Angrist & Pischke, *supra* note 16, at 113–218.

[19] *See* William R. Shadish, Thomas D. Cook, & Donald T. Campbell, Experimental and Quasi-Experimental Design for Generalized Causal Inference (2002).

[20] *See* Law, *supra* note 9, at 390–5.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

of observational studies and experiments. Furthermore, it is also possible to combine quantitative approaches with qualitative case study analyses.[21] The quantitative analysis could then establish the existence of a correlation, while the case studies could rule out the influence of non-observed omitted variables or establish the direction of causality. This is not the place to address these different techniques in detail. Nevertheless, the preceding discussion has shown that there is no one-size-fits-all approach or statistical all-purpose weapon that can be used to address every desired question. Instead, researchers need a high amount of creativity to come up with an adequate research design for a specific research question. This also means that there are questions we might not be able to address at all with the methodological instruments currently at our disposal.

3. Institutional design

Broadly speaking, empirical research on the institutional design of constitutions can take two directions. The research question can address either the consequences of a design choice or the reasons for this choice. While in the first case the research design is concerned with the effects of a constitutional arrangement, in the second case it is concerned with the question of why a certain arrangement has come about. It comes as no surprise that these research interests tend to correspond with certain disciplinary backgrounds. While economists tend to be interested in the effects of legal rules, it is rather the political scientists and legal scholars who are interested in explaining the reasons for certain institutional arrangements.[22] Research in both areas faces significant methodological challenges. The challenges concern the aforementioned omitted-variable bias, but also problems of conceptualization and measurement. The following discussion will move from a selection of interesting studies and results to their shared problems and challenges.

3.1. Consequences of design choices

It is intuitive that, as constitutional design choices matter for political processes, they might also matter for economic outcomes. Such economic effects were studied in the seminal work of Torsten Persson and Guido Tabellini.[23] Persson and Tabellini place great emphasis on studying the effects of two categorical features of constitutions: the forms of government, and the electoral rules. Their results indicate that both presidential forms of government and majoritarian electoral rules have a negative effect on government spending and the size of welfare state programs.[24] On the other hand, they find no robust significant effect on the level of corruption or productivity.[25]

21 *See, e.g.*, Zachary Elkins, Tom Ginsburg, & James Melton, The Endurance of National Constitutions (2009).

22 On the economists' perspective, see the survey articles by Voigt, *Empirical Constitutional Economics: Onward and Upward?*, *supra* note 11, at 11; Stefan Voigt, *Positive Constitutional Economics II: A Survey of Recent Developments*, 146 Pub.Choice 205 (2011). Voigt himself is pleading for a broader perspective.

23 Torsten Persson & Guido Tabellini, The Economic Effects of Consitutions (2003).

24 *Id.*, at 155–86.

25 *Id.*, at 187–217.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

The work by Persson and Tabellini also stands out due to its methodological approach, as it applies rigorous econometric methods to matters of constitutional design.

From the perspective of comparative law, the monograph by Zachary Elkins, Tom Ginsburg, and James Melton[26] on the endurance of constitutions is widely considered as the starting point for rigorous, quantitative empirical work. Based on theoretical considerations, Elkins and colleagues ask which design choices benefit the endurance of a constitution. Thus, they analyze the consequences of institutional design for the institution itself. They use data from the Comparative Constitutions Project (CCP) on all written constitutions from 1789 to 2005 and identify three design variables that influence a constitution's longevity. First, the longevity of the constitution benefits from a strong inclusivity of the constitution-making process.[27] Second, endurance is affected by a constitution's flexibility; however, the relationship between these variables is not linear: instead, the curve is inversely u-shaped.[28] This means that both types of amendment procedures—the too flexible one as well as the one that is too difficult—hurt constitutional longevity; therefore, constitutional designers have the task of finding the right balance. Third, a constitution is more likely to endure the more specific it is.[29] According to the authors, a higher level of detail indicates that elites invested in solving their conflicts with specified compromises during the constitution-making process.[30]

Both studies are not only landmark studies for the empirical research on constitutions, but they also highlight the challenges for this research. The challenges concern, in particular, the identification of causality.[31] The most important challenge for the identification of causality in the discussed studies is a potential omitted-variable bias. Even if the studies observe interesting correlations, these correlations could well be due to unobserved variables that influence both the explanatory and the explained variables. Persson and Tabellini as well as Elkins, Ginsburg, and Melton employ different strategies to address the omitted-variable bias.

In the case of Persson and Tabellini's study, there might be an unobserved variable that influences both the adoption of a specific political system and the size of the welfare state. It is not implausible that a rather collectivist culture leads to societal

26 Elkins, Ginsburg, & Melton, *supra* note 21.

27 *Id.*, at 139–40.

28 *Id.*, at 140–41.

29 *Id.*, at 141. *See further* Mila Versteeg & Emily Zackin, *Constitutions Unentrenched: Toward an Alternative Theory of Constitutional Design*, 110 Am. Pol. Sci. Rev. 657 (2016) (observing that there is a trend in younger constitutions to more specificity).

30 Elkins, Ginsburg, & Melton, *supra* note 21, at 103.

31 There are also some problems of conceptualization and measurement. Ginsburg and Melton discuss the issue of measuring the flexibility of a constitution themselves in a later contribution; *see* Tom Ginsburg & James Melton, *Does the Constitutional Amendment Rule Matter At All? Amendment Cultures and the Challenges of Measuring Amendment Difficulty*, 13 Int'l J. Const. L. 686 (2015). For a critique of the conceptualization and measurement of the specificity variable, see Samuel Issacharoff, *Measuring Law: Version 2.0* (May 1, 2020) (unpublished manuscript) (on file with the authors) (arguing that counting the words of constitutions has only limited meaning). However, as these measurement problems are not likely to lead to systematic biases, we will not discuss this issue further.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

preferences for egalitarianism and political compromise. This might, in turn, result in the adoption of a representative electoral system as well as translate into higher welfare payments. In such a scenario, both the representative electoral system and the bigger welfare state would thus only be an indication of a more collectivist culture, so that a correlation between the two does not necessarily indicate a causal relationship.

Persson and Tabellini employ different strategies to address this challenge. The most important strategy is an instrumental variables estimation.[32] As outlined above, the success of an instrumental-variables strategy depends on the instrument being a variable that influences the dependent variable only through the explanatory variable.[33] In the concrete case, the instrument should thus only influence economic performance, the size of the budget, and the welfare state through the electoral system and the degree of presidentialism, but it should not have an independent influence on the dependent variable or any of the omitted variables.

Persson and Tabellini use three main instruments—the date of the adoption of the constitution, the distance of the country from the equator, and the share of the population speaking European languages. As Daron Acemoglu has shown in a detailed review of the book, none of these instruments is convincing.[34] The date of adoption might have an influence on the shape of the constitution without influencing any of the dependent variables. However, the influence of the instrument is too weak to carry the analysis.[35] The distance from the equator and the fraction of the population speaking European languages are popular instruments for the quality of political institutions in the comparative politics literature.[36] The intuition behind these instruments is that strong institutions in modern times originated in Europe.[37] The further a country is from the equator, the more appealing was its climate for European settlements and thus for the introduction of European institutions. Similarly, the share of European languages is also an indicator for European settlements. One may discuss whether these two variables are indeed good proxies for the quality of institutions in general. However, it is implausible that they are also an explanation for the shape of specific institutions, such as presidentialism or the electoral system.[38] Moreover, they do not really explain the development of institutions within Europe, even though European countries are also part of the analyzed sample.[39]

A replication study underscores the methodological difficulties: Lorenz Blume and colleagues tried to replicate the results with an expanded dataset that now contains 115 countries instead of eighty-five. While they were able to replicate the results

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

32 Persson & Tabellini, *supra* note 23, at 113–53.

33 *See* Section 1.2.b.

34 Acemoglu, *supra* note 14, at 1034–43.

35 *Id.* at 1035–6.

36 *Seminally* Robert E. Hall & Charles I. Jones, *Why Do Some Countries Produce So Much More Output Per Worker Than Others?*, 114 Q. J. Econ. 83 (1999).

37 *Id.* at 100–2.

38 Acemoglu, *supra* note 14, at 1041–3.

39 *Id.* at 1040.

concerning the effects of majoritarian electoral rules, they could not do so with the effects of presidential forms of government.[40] The bottom line is that Persson and Tabellini do indeed show interesting correlations but the establishment of causal links, however, seems to require at least further research.

The study of Elkins, Ginsburg, and Melton also faces the challenge of controlling for potential confounding factors. There might be variables that may well influence both a design feature of a constitution as well as a constitution's longevity. If, for instance, the specificity of a constitution does indeed indicate that political elites settled their conflicts, as the authors claim,[41] then the constitution's longevity might not depend on its specificity, but much rather on the continuing cooperation of the political elites. Specificity would then be an indication for elite cooperation, but not a factor that has a causal effect on the constitution's endurance.

Statistically, Elkins, Ginsburg, and Melton try to avoid spurious correlations by including control variables in their regression analysis.[42] Of course, they are fully aware that the control variables might not capture all omitted variables. Thus, they complement their quantitative results with qualitative case studies, in order to examine the assumed causal mechanisms more closely.[43] These case studies provide depth where the quantitative approach offers breadth. This strategy turns out to be at least partly successful: There remains a lack of qualitative evidence for the stabilizing function of a constitution's specificity and the fear that an omitted variable might be the driving force cannot be refuted. But, at least, the important role of inclusivity in the constitution-making process is reinforced—it seems that in case of more inclusive processes, new governments are less tempted to replace an old constitution to mark their dominance.[44]

3.2. Reasons for design choices

Empirical research on constitutional design does not only look at the effect of design choice, but also at the reasons for such choices. The latter discussion has touched on a wide array of issues:[45] Empirical studies range from examining the reasons

[40] Lorenz Blume, Jens Müller, Stefan Voigt, & Carsten Wolf, *The Economic Effects of Constitutions: Replicating – and Extending – Persson and Tabellini*, 139 Public Choice 197 (2009).

[41] Elkins, Ginsburg, & Melton, *supra* note 21, at 103.

[42] *Id.* at 111–21.

[43] *Id.* at 147–206.

[44] *Id.* at 188.

[45] Next to the issues mentioned explicitly in the following, there are also numerous studies on other issues regarding constitutional design, see, e.g., James Melton, Zachary Elkins, Tom Ginsburg, & Kalev Leetaru, *On the Interpretability of Law: Lessons From the Decoding of National Constitutions*, 43 B. J. Pol. Sci. 399 (2012); Tom Ginsburg, Zachary Elkins, & James Melton, *Do Executive Term Limits Cause Constitutional Crises?*, *in* Comparative Constitutional Design 350 (Tom Ginsburg ed., 2012); Rosalind Dixon & Richard Holden, *Constitutional Amendment Rules*, *in* Comparative Constitutional Design, *supra*, at 195; Dawood I. Ahmed & Tom Ginsburg, *Constitutional Islamization and Human Rights: The Surprising Origin and Spread of Islamic Supremacy in Constitutions*, 54 Va. J. Int'l L. 615 (2014); Ginsburg & Melton, *supra* note 31; Versteeg & Zackin, *supra* note 29.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

for constitutional entrenchment clauses[46] over the determinants for emergency provisions[47] to judicial independence[48] and the constitutional openness for international law.[49] In this section, we would like to focus on a study by Tom Ginsburg and Mila Versteeg, which examines the reasons for why states introduce constitutional review.[50]

Constitutional review is not only subject to a controversial normative discussion on the legitimacy of courts controlling the legislature.[51] Its introduction to a political system also poses an empirical puzzle, as legislative majorities and governments have little incentive to put themselves under judicial scrutiny. In their paper, Ginsburg and Versteeg want to test different theories of why political elites might want to introduce constitutional review.[52] They differentiate between four different explanations for the introduction of constitutional review—the resolution of competency issues in federal states; ideational reasons, such as the desire to protect the rule of law; strategic reasons, according to which constitutional review is an insurance policy in case the government loses power; and diffusion, i.e. that states adopt constitutional review as a response to constitutional developments in other states.[53] In their empirical analysis, they only find robust support for the notion of constitutional review as an insurance

[46] Michael Hein, *Impeding Constitutional Amendments: Why Are Entrenchment Clauses Codified in Contemporary Constitutions?*, 53 Acta Politica 1 (2018).

[47] Christian Bjørnskov & Stefan Voigt, *The Architecture of Emergency Constitutions*, 16 Int'l J. Const. L. 101 (2018).

[48] *See* Lars P. Feld & Stefan Voigt, *Economic Growth and Judicial Independence: Cross-Country Evidence Using a New Set of Indicators*, 19 Eur. J. Pol. Econ. 497 (2003); Bernd Hayo & Stefan Voigt, *Explaining* De Facto *Judicial Independence*, 27 Int'l Rev. L. & Econ. 269 (2007); Julio Ríos-Figueroa & Jeffrey K. Staton, *An Evaluation of Cross-National Measures of Judicial Independence*, 30 J. L. Econ. & Org. 104 (2014); James Melton & Tom Ginsburg, *Does De Jure Judicial Independence Really Matter?*, 2 J. L. & Cts. 187 (2014); Bernd Hayo & Stefan Voigt, *Explaining Constitutional Change: The Case of Judicial Independence*, 48 Int'l Rev. L. & Econ. 1 (2016).

[49] Tom Ginsburg, Svitlana Chernykh, & Zachary Elkins, *Commitment and Diffusion: How and Why National Constitutions Incorporate International Law*, 2008 U. Ill. L. Rev. 201 (2008).

[50] Tom Ginsburg & Mila Versteeg, *Why Do Countries Adopt Constitutional Review?*, 30 J. L. Econ. & Org. 587 (2014).

[51] *See, e.g.*, Alexander M. Bickel, The Least Dangerous Branch. The Supreme Court at the Bar of Politics (1962); Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 Ind. L.J. 1 (1971); William H. Rehnquist, *The Notion of a Living Constitution*, 54 Tex. L. Rev. 693 (1976); John Hart Ely, Democracy and Distrust: A Theory of Judicial Review (1980); Bruce Ackerman, *The Storrs Lectures: Discovering the Constitution*, 93 Yale L.J. 1013 (1984); Antonin Scalia, *Originalism: The Lesser Evil*, 57 U. Cincinnati L. Rev. 849 (1989); Bruce Ackerman, *Constitutional Politics/constitutional Law*, 99 Yale L.J. 453 (1989); Robert H. Bork, The Tempting of America: The Political Seduction of the Law (1990); Bruce Ackerman, We the People: Foundations (1991); Jeremy Waldron, *A Right-Based Critique of Constitutional Rights*, 13 Oxford J. Legal Stud. 18 (1993); Ronald Dworkin, Freedom's Law: The Moral Reading of the American Constitution (1996); Jeremy Waldron, Law and Disagreement (1999); Mark Tushnet, Taking the Constitution Away From the Courts (1999); Larry D. Kramer, The People Themselves: Popular Constitutionalism and Judicial Review (2004); Jeremy Waldron, *The Core of the Case Against Judicial Review*, 115 Yale L.J. 1346 (2006); Dimitris Kyritsis, *Representation and Waldron's Objection to Judicial Review*, 26 Oxford J. Legal Stud. 733 (2006); Richard Bellamy, Political Constitutionalism: A Republican Defense of the Constitutionality of Democracy (2007); Michael J. Perry, Constitutional Rights, Moral Controversy, and the Supreme Court (2009).

[52] Ginsburg & Versteeg, *supra* note 50.

[53] *Id.* at 588.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

policy against the risk of losing power.[54] By contrast, they do not find support for any of the other hypotheses.

The idea that constitutional review serves as an insurance policy for the government in cases in which the political system is not dominated by a single political group or party is theoretically plausible.[55] But the empirical analysis is not without problems. These concern both the conceptualization and measurement of the analyzed data, as well as the statistical techniques that are used to address a potential omitted-variable bias.

a) Conceptualization and measurement

Most of the independent variables that the authors use in their analysis to explain the introduction of constitutional review are latent variables, which cannot be directly observed. Therefore, Ginsburg and Versteeg have to rely on proxies, which measure the explanatory concepts indirectly. The reliability of these proxies as representations of the analyzed concepts depends on the extent to which they are indeed correlated with the latter. If there is only a weak correlation between the directly measured variables and the underlying concepts, this leads to a considerable measurement error and in turn to biased results.[56]

Crucially, Ginsburg and Versteeg build on the political insurance variable, which is supposed to capture the extent to which the ruling party fears losing power. In order to measure this latent variable indirectly, the authors use the difference between the share of seats of the largest party in the lower house of parliament and the share of seats of the second-largest party.[57] For the case of a parliamentary democracy, this seems to be a reasonable choice. However, this choice does not account for other institutional arrangements. In a presidential or semi-presidential democracy, the mere difference between the two largest parties in parliament gives us little information about the confidence of the government to stay in power without additional information on whether the president is a partisan of the strongest party in parliament.

It is even less conclusive as a proxy for the threat to the ruling elites in an authoritarian régime. In a one-party state, the composition of parliament may not change. Yet, the grip on power of the government may well do so. In such cases, changes in political power are not reflected in parliament's composition, and thus are not in the least captured by the measure that the authors propose. Furthermore, we may have authoritarian régimes, in which a modest party competition is tolerated. In such cases, the more useful information would probably be whether the largest party in parliament is aligned with the president or not. Take Iran as an example. An electoral victory of the moderate Moderation and Development Party may well signal discontent with the Supreme Leader. However, this is not captured by merely measuring the

[54] *Id.* at 606–13.

[55] *See, e.g.*, Samuel Issacharoff, *Constitutional Courts and Democratic Hedging*, 99 Geo. L.J. 961 (2011); Samuel Issacharoff, Fragile Democracies: Contested Power in the Era of Constitutional Courts (2015).

[56] *See* Kenneth A. Bollen, Structural Equations With Latent Variables 154–67 (1989).

[57] Ginsburg & Versteeg, *supra* note 50, at 604.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

difference in the share of seats of the two largest parties. And, on the contrary, a landslide victory of the opposition may be a much stronger signal of discontent than a narrowly eked out majority.

Another explanatory concept that Ginsburg and Versteeg try to measure indirectly is the attitude towards the rule of law. They do so in order to test whether a more positive attitude towards the rule of law leads to a higher likelihood of adopting constitutional review. In order to measure the attitude towards the rule of law, they rely on two variables—the number of legal publications per capita and the number of law schools per capita.[58] While it is undeniable that these two variables are probably correlated with a legalistic culture, the question remains how strong this correlation is. The number of legal publications and the number of law schools presumably also depend on other factors than mere fondness for the rule of law. They are probably highly correlated with economic development: a rich country may have more law schools and more scholars churning out legal publications even if its relationship with the rule of law may be ambiguous.

Finally, Ginsburg and Versteeg employ diffusion as a latent explanatory variable. However, there is a gap between the diffusion channels that they describe in their theoretical part and the suggested empirical measurement of diffusion. Theoretically, they identify four different diffusion channels—coercion, competition, learning, and acculturation.[59] But not all of these four channels are adequately captured by the variables that they include in their analysis—shared common colonizer, shared dominant religion, shared language, and geographic proximity.[60] While the shared common colonizer and geographic proximity may at least partly capture coercion and competition, this is much less clear for shared language and religion as proxies for learning and acculturation. It seems much more plausible that states introducing constitutional review rely on constitutions with considerable authority, regardless of whether there is a shared language or religion.[61]

b) Addressing the omitted-variable bias

Ginsburg and Versteeg dedicate surprisingly little room to discussing a potential omitted-variable bias.[62] The main fix that they use in their basic model is a set of country fixed effects that is supposed to control for country characteristics.[63] Nonetheless, it is not implausible that both the explanatory variables and the dependent variable might be influenced by a time-variant third factor.

[58] *Id.* at 603.

[59] *Id.* at 596–7.

[60] *See id.* at 605–6.

[61] This is particularly the case for matters of organization. In the case of culturally sensitive matters, such as fundamental rights catalogues, religion might play a more important role. It is quite telling in this respect that the authors themselves refer to the case of the Irish Free State Constitution from 1922, in which the drafters translated other pre-existing constitutions in order to learn from them, which indicates that *shared* language is probably not a strong predictor for the learning diffusion mechanism, *see id.* at 597.

[62] *See id.* at 616.

[63] *Id.* at 598–9.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

Arguably, constitutional review is often introduced in transitional phases. When countries draft a new constitution or implement a major constitutional reform, they might seize this opportunity to introduce a constitutional review mechanism. However, this constitutional transition might lead not only to the introduction of constitutional review, but also to a major shift in the composition of parliament. An obvious case would be the transition from an authoritarian one-party state to a parliamentary democracy, which also comes with the introduction of constitutional review. In such a case, we would see that the introduction of constitutional review is accompanied by a considerable decrease in the political insurance variable. However, constitutional review is not necessarily introduced as an insurance policy, but may be based on many other reasons.

Ginsburg and Versteeg try to address this challenge by adding a democracy indicator from the Polity IV dataset into their robustness checks.[64] However, there may be constitutional reforms that do not influence the democracy score, and, *vice versa*, the level of democracy may change without constitutional reform. One may only think of an authoritarian one-party state introducing controlled parliamentary elections and judicial review as window-dressing in order to appease influential trade partners. In such a case, we would have a constitutional reform that is accompanied by the introduction of constitutional review without a change in the democracy score because the quality of democracy has not changed in practice.

The critique highlights some of the potential pitfalls in analyzing the reasons for constitutional design choices. We do not suggest that constitutional review does not have the function of an insurance policy in cases where there is no single dominant political elite. Indeed, the explanation is highly plausible. Nevertheless, we wanted to point out that the empirical evidence is weaker than Ginsburg and Versteeg suggest in their paper. More importantly, the latter does not provide evidence that there might not be other reasons for the introduction of constitutional review.[65] On the one hand, we saw that the measured variables might only be weakly correlated with the analyzed concepts. On the other hand, the lack of statistical significance does not prove the absence of a causal effect. Instead, it may also be due to the lack of statistical power.

4. Constitutional rights

The empirical research on constitutions is not limited to the study of institutional design. While not quite as extensive, there is a growing body of research on constitutional rights. These studies can be put into two broad categories. First, there is a lively and controversial discussion about the diffusion of constitutional rights (Section 4.1). Second, there is also research addressing the effectiveness of constitutional rights (Section 4.2).

[64] *Id.* at 616.

[65] But see *id.* at 617 (underlining the importance of their finding that diffusion-based theories lack empirical support).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

4.1. Diffusion of constitutional rights

An important strand of the empirical comparative constitutional law literature has focused on the diffusion of constitutional rights: To what extent do existing constitutional rights catalogues influence the shape of rights sections in new constitutions? Which existing rights catalogues are chosen as models and why are they chosen?

A study by David Law and Mila Versteeg on the evolution of global patterns of constitutional rights makes mostly descriptive claims.[66] The authors try to "map […] the global constitutional landscape" by estimating ideal points for each constitutional rights catalogue.[67] They argue that the position of a constitution on this map can largely be explained by two variables: the comprehensiveness of the constitution and its ideology, that is, the question whether it is a liberal or a statist constitution.[68] Over time, the authors observe both a convergence and a polarization. They identify two ideological categories and argue that constitutions tend to converge within their ideological category, but that these two categories are increasingly drifting apart from each other.[69]

Benedikt Goderis and Mila Versteeg analyze the diffusion of constitutional rights and examine why states include specific rights in their constitutions.[70] They find some interesting correlations: the adoption of a constitutional right is correlated with the adoption of the same right by the former colonizer, by countries with the same legal origin and/or religion, and by countries with the same former colonizer or the same aid donor.

A controversial discussion focuses on the question of which constitutional texts or international documents are most influential for the shape of subsequent constitutional rights catalogues. In a high-profile study,[71] David Law and Mila Versteeg observe that the influence of the US Constitution has continuously decreased.[72] They also argue that no other rights catalogue has taken its place. The Canadian Charter of Rights and Freedoms was influential in the Commonwealth, but not beyond.[73] The constitutions of Germany, India, and South Africa might be popular reference points in the comparative constitutional law literature, but they have little influence on the textual shape of new constitutions.[74] Finally, the authors state that the Universal Declaration of Human Rights (UDHR) or international human rights treaties did not have a significant influence on modern constitution-making.[75] The latter result is

[66] David S. Law & Mila Versteeg, *The Evolution and Ideology of Global Constitutionalism*, 99 Cal. L. Rev. 1163 (2011).

[67] *Id.* at 1204.

[68] *Id.* at 1217–6.

[69] *Id.* at 1233–43.

[70] Benedikt Goderis & Mila Versteeg, *The Diffusion of Constitutional Rights*, 39 Int'l Rev. L. & Econ. 1 (2014).

[71] *See* Adam Liptak, *"We the People" Loses Appeal with People around the World*, N.Y. Times, (Feb. 6, 2012), https://nyti.ms/3vj1Lwv.

[72] David S. Law & Mila Versteeg, *The Declining Influence of the United States Constitution*, 87 N.Y.U. L. Rev. 762, 779–809 (2012).

[73] *Id.* at 809–23.

[74] *Id.* at 823–33.

[75] *Id.* at 833–50.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

contested by a study of Zachary Elkins, Tom Ginsburg, and Beth Simmons, who find a significant convergence of human rights catalogues around the globe and argue that the UDHR and the International Covenant on Civil and Political Rights (ICCPR) play a significant coordinating role in this convergence.[76]

Studies on the diffusion of constitutional rights face two kinds of challenges, which we would like to analyze in further detail in the following two sections. The first challenge concerns the conceptualization and measurement of constitutional rights that are contained in a constitution. While this might sound straightforward prima facie, the devil lies in the detail (Section 4.1(a)). To the extent that studies on diffusion not only observe correlations and make descriptive claims, but also identify causal trends, the identification strategy used for establishing causality poses a second challenge (Section 4.1(b)).

a) Problems of conceptualization, operationalization, and measurement

Prima facie, conceptualizing, operationalizing, and measuring whether a constitution contains a specific right or not seems straightforward. However, researchers have to make certain choices in this process. While these choices probably do not bias the results of the analysis, they do influence their interpretation. Consequently, many claims made by the analyzed studies are less broad and bold than they seem at first sight. In the following, we would like to highlight three problems regarding the conceptualization of variables and how these influence the interpretation of the results.

First, researchers examining diffusion have to determine the breadth of the categories that they want to compare. For example, the protection of personal liberty comes in different forms. Some constitutions protect the right to liberty in general, while others prohibit arbitrary arrest and detention. Whether to code them as the same right or as different rights depends on the research question. If you want to measure similarity, it seems preferable to construct broad categories,[77] while the construction of separate categories seems preferable if you want to examine influence and assume that influence is also reflected in the wording of a constitutional provision.[78]

Second, and more importantly, researchers have to make a choice about what counts as a constitutional right. The problem arises in particular if the constitution incorporates the texts of international human rights treaties: Should the full content of these treaties be regarded as part of the domestic constitution? David Law and Mila Versteeg made a distinction in this respect: they only counted the provisions of an international human rights instrument as part of the constitution if they were specifically enumerated in the constitution itself.[79] The European Convention on Human

[76] Zachary Elkins, Tom Ginsburg, & Beth A. Simmons, *Getting to Rights: Treaty Ratification, Constitutional Convergence, and Human Rights Practice*, 54 Harv. Int'l L.J. 61, 74–88 (2013). *See also* G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948); International Covenant on Civil and Political Rights, Dec. 16, 1966, 999 U.N.T.S. 171.

[77] In this vein, Law & Versteeg, *supra* note 72, at 1191.

[78] *See* Goderis & Versteeg, *supra* note 70, at 4–5. The latter is no necessary conclusion, see below, Section 4.1(b).

[79] Law & Versteeg, *supra* note 72, at 1189.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

Rights (ECHR) was thus considered to be an integral part of the UK Constitution because its provisions were set forth in full in the appendix of the UK Human Rights Act.[80] By contrast, the ECHR was not considered to be part of the Austrian Constitution, because the latter only makes a general and dynamic reference to the ECHR without listing the specific rights.[81]

The basic intuition behind this choice is probably that general references to human rights treaties come in various forms and colors. Sometimes, they form part of the preamble;[82] sometimes, they are obligations to interpret rights in the light of international human rights documents;[83] and sometimes, international law generally has constitutional or supra-constitutional status. It is often difficult to determine the precise status of international human rights within the domestic constitutional order without having a detailed look at the constitutional jurisprudence. Nevertheless, the distinction between the Austrian and the UK example seems arbitrary. There is probably no stronger form of influence of an international human rights document if the constitution directly refers to it and incorporates it even if the specific provisions are not explicitly listed. While this is a legitimate conceptual choice, it limits the strength of the claim that Law and Versteeg are making. By not counting general references to human rights treaties, Law and Versteeg adopt an interpretation of influence that underestimates the real impact of international human rights documents on the constitutional practice.

Finally, all analyzed studies limit themselves to rights that are explicitly contained in the constitutional text.[84] While this is, again, a legitimate conceptual choice, it probably misses the most important channel of diffusion that is currently discussed in the comparative law literature—diffusion through courts.[85] Notably, it is impossible to appreciate the influence of the ECHR on domestic constitutional orders in

[80] *Id. See also* Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, 213 U.N.T.S. 222; Human Rights Act, c 42, 1998 (U.K.).

[81] *See* Bundesverfassungsgesetz [B-VG] [Federal Constitutional Statute], Mar. 4, 1964, 1964 Bundesgesetzblatt für die Republik Österreich [BGBl.] 623, art. II Z. 7 (Austria), www.ris.bka.gv.at/Dokumente/BgblPdf/1964_59_0/1964_59_0.pdf.

[82] See the example of Congo Const., Jan. 20, 2002, cited by Law & Versteeg, *supra* note 72, at 1189.

[83] *See, e.g.*, S. Afr. Const. § 39 (1)(b), 1996.

[84] Law & Versteeg, *supra* note 72, at 1188; Goderis & Versteeg, *supra* note 70, at 4; Elkins, Ginsburg, & Simmons, *supra* note 76, at 69.

[85] See only the vast literature on judicial dialogue between constitutional courts, e.g. Anne-Marie Slaughter, *A Typology of Transjudicial Communication*, 29 U. Rich. L. Rev. 99 (1994); Christopher McCrudden, *A Common Law of Human Rights?: Transnational Judicial Conversations on Constitutional Rights*, 20 Oxford J. Leg. Stud. 499 (2000); Bijon Roy, *An Empirical Survey of Foreign Jurisprudence and International Instruments in Charter Litigation*, 62 U. Toronto Faculty L. Rev. 99 (2004); Eyal Benvenisti, *Reclaiming Democracy: The Strategic Uses of Foreign and International Law By National Courts*, 102 Am. J. Int'l L. 241 (2008); David S. Law & Wen-Chen Chang, *The Limits of Global Judicial Dialogue*, 86 Wash. L. Rev. 523 (2011); Elaine Mak, *Why Do Dutch and UK Judges Cite Foreign Law?*, 70 Cambridge L.J. 420 (2011); Anne Meuwese & Marnix Snel, *"Constitutional Dialogue": An Overview*, 9 Utrecht L. Rev. 123 (2013); Elaine Mak, *Globalisation of the National Judiciary and the Dutch Constitution*, 9 Utrecht L. Rev. 36 (2013); Ryan C. Black, Ryan J. Owens, & Jennifer L. Brookhart, *We Are the World: The U.S. Supreme Court's Use of Foreign Sources of Law*, 46 B.J. Pol. Sci. 891 (2014).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

Europe without analyzing the case law of domestic apex courts regarding the status of the Convention within the domestic legal order. For example, the right to a fair trial, enshrined in Article 6 ECHR, does not have a counterpart in the German Basic Law. Nevertheless, it is, de facto, protected in the same way in the German constitutional order because the German Federal Constitutional Court considers many of its individual elements to be part of the rule-of-law guarantee in Article 20 of the German Basic Law.[86]

Consequently, the apparent punchline of Law and Versteeg's study on the declining influence of the US Constitution is far less steep than it appears at first sight. When the authors refer to the US Constitution, then they mean the text of the constitution, but not US constitutionalism. However, if we think of the influence of US constitutional law on other constitutional orders, we would rather refer to the case of law of the US Supreme Court[87] or US constitutional theory. Similarly, the jurisprudence of the German Federal Constitutional Court seems to be far more influential internationally than the text of the German Basic Law.

b) Identifying causality

Most of the studies discussed in this part do not make causal claims. Law and Versteeg's study maps the global constitutional landscape and is therefore rather more engaged in a categorization than in the identification of causal mechanisms. Similarly, Goderis and Versteeg's study on diffusion stops short of making causal claims, instead only observing correlations. However, the studies on the influence of particular constitutional rights documents on subsequent rights try to identify causality. If we say that one text influences another, then we claim that the shape of the former text has a causal effect on the shape of the subsequent text.

In an elegant critique, Zachary Elkins, Tom Ginsburg, and James Melton argued that Law and Versteeg's study on the declining influence of the US Constitution does not actually measure influence.[88] Law and Versteeg try to identify influence by measuring similarity.[89] However, influence is not the same as similarity.[90] Two documents might

[86] *See, e.g.*, Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], Mar. 11, 1975, 39 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 156, 163; Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], May 26, 1981, 57 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 250, 274; Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], Jan. 24, 2001, 103 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 44, 64; Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], May 4, 2004, 110 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 339, 342; Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], Jan. 15, 2009, 122 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 248, 271–2; Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court], Dec. 7, 2011, 130 Entscheidungen des Bundesverfassungsgerichts [BVerfGE] 1, 25 (Ger.). *See also* Grundgesetz [GG] [Basic Law], *translation at* www.gesetze-im-internet.de/englisch_gg/index.html.

[87] *See, similarly*, Vicki C. Jackson, *Comment on Law and Versteeg*, 87 N.Y.U. L. Rev. 2102, 2106 (2012).

[88] Zachary Elkins, Tom Ginsburg, & James Melton, *Comments on Law and Versteeg's the Declining Influence of the United States Constitution*, 87 N.Y.U. L. Rev. 2088 (2012).

[89] Law & Versteeg, *supra* note 72, at 779–81.

[90] Elkins, Ginsburg, & Melton, *supra* note 88, at 2093–4.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

be similar without being influenced by each other, while influence does not necessarily lead to similarity. It is not implausible that the shape of rights has changed over time from the prohibition of certain concrete actions to more abstract rights and liberties. For example, the Magna Carta or the Fourth Amendment to the US Constitution contain prohibitions of unlawful arrest. By contrast, modern human rights documents often tend to guarantee a general right to liberty without necessarily specifying concrete prohibited actions.[91] However, does this mean that the former have not at all influenced the latter? Or does it only signify that the form of a specific right has evolved over time? If we draw the latter conclusion, then dissimilarity does not automatically allow the conclusion of a lack of influence.

By contrast, Elkins, Ginsburg, and Simmons rely on a difference-in-difference analysis, in which they track the inclusion of specific rights in constitutional documents before and after the adoption of the UDHR and the ICCPR.[92] In particular, they analyze whether the prevalence of rights included in these international documents has grown faster than the prevalence of rights not included.[93] This method is not waterproof either. It cannot exclude the effect of potential unobserved variables, that is, factors that have influenced the inclusion of a specific right in the UDHR/ICCPR and at the same time in subsequent domestic constitutions. Nevertheless, there are, prima facie, no obvious candidates for such a confounding variable, so that the analysis has at least a significant plausibility.

4.2. Effectiveness of constitutional rights

The effectiveness of individual rights has, for a long time, been subject to controversial debate. While legal scholars began relatively early to investigate the effectiveness of international human rights,[94] the effectiveness of constitutional rights has predominantly been studied by political scientists.[95] Nevertheless, the effect of constitutional rights on government conduct has recently also been addressed by legal scholars. In particular, Adam Chilton and Mila Versteeg have analyzed the effectiveness of

[91] Coded as separate guarantees by Law & Versteeg, *supra* note 72, at 1188–9.

[92] Elkins, Ginsburg, & Simmons, *supra* note 76.

[93] Elkins, Ginsburg, & Simmons, *supra* note 76, at 78–9.

[94] *See*, e.g., Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 Yale L.J. 1935 (2002); Ryan Goodman & Derek Jinks, *Measuring the Effects of Human Rights Treaties*, 14 Eur. J. Int'l L. 171 (2003); Eric Neumayer, *Do International Human Rights Treaties Improve Respect for Human Rights?*, 49 J. Conflict Resolution 925 (2005); Beth A. Simmons, Mobilizing for Human Rights: International Law in Domestic Politics (2009); Richard A. Nielsen & Beth A. Simmons, *Rewards for Ratification: Payoffs for Participating in the International Human Rights Regime?*, 59 Int'l Stud. Q. 197 (2015).

[95] *See*, e.g., Kathleen Pritchard, *Comparative Human Rights: An Integrative Explanation*, 13 S. Afr. J. Pol. Stud. 24 (1986); Christian A. Davenport, *"Constitutional Promises" and Repressive Reality: A Cross-National Time-Series Investigation of Why Political and Civil Liberties Are Suppressed*, 58 J. Pol. 627 (1996); Frank Cross, *The Relevance of Law in Human Rights Protection*, 19 Int'l Rev. L. & Econ. 87 (1999); Linda Kemp Keith, *Constitutional Provisions for Individual Rights (1977–1996): Are They More Than Mere "Window Dressing?,"* 55 Pol. Res. Q. 111 (2002); Linda Kemp Keith, C. Neal Tate, & Steven C. Poe, *Is the Law a Mere Parchment Barrier to Human Rights Abuse?*, 71 J. Pol. 644 (2009); Jonathan Fox & Deborah Flores, *Religions, Constitutions, and the State: A Cross-National Study*, 71 J. Pol. 1499 (2009); James Melton, Do Constitutional Rights Matter? (2014).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

constitutional rights and their implementation by constitutional courts in several studies.[96] Most notably, they have studied and compared the effectiveness of organizational rights and individual rights in constitutions.[97] They argue that the inclusion of organizational rights, such as the right to form political parties, freedom of association, or the right to unionize, has a positive effect on their respect in practice.[98] By contrast, traditional individual rights, such as freedom of expression or movement, do not make a difference.[99]

In further studies, the two authors have extended their analysis of the effectiveness of constitutional rights to other fields. They contend that the inclusion of torture prohibitions in constitutions does not have a significant effect on the practice of torture,[100] that social rights do not have a measurable influence on parliamentary budget decisions,[101] and that the ability of constitutional courts to protect constitutional rights is limited.[102] Such studies on the effectiveness of individual rights face two challenges. On the one hand, they have to conceptualize, operationalize, and measure the dependent and the independent variable (Section 4.2(a)). On the other hand, they have to accommodate potential selection effects in their identification strategy when trying to establish causality (Section 4.2(b)).

a) Problems of conceptualization, operationalization, and measurement

The problem of conceptualization, operationalization, and measurement concerns both the independent and the dependent variable. Regarding the independent variable, Chilton and Versteeg—as the studies on diffusion—only measure the existence of rights which are explicitly mentioned in the constitutional text.[103] While this was still a legitimate conceptual choice when studying the diffusion of constitutional rights,[104] the exclusive focus on rights which are expressly mentioned in the constitution may bias the results when analyzing the effectiveness of constitutional guarantees.

As mentioned, the coding of the Austrian Constitution did not take into account that the constitution has explicitly incorporated the ECHR so that the Convention guarantees have, in practice, the same status as constitutional rights in the Austrian

[96] Adam Chilton & Mila Versteeg, *The Failure of Constitutional Torture Prohibitions*, 44 J. Legal Stud. 417 (2015); Adam Chilton & Mila Versteeg, *Do Constitutional Rights Make a Difference?*, 60 Am. J. Pol. Sci. 575 (2016); Adam Chilton & Mila Versteeg, *Rights Without Resources: The Impact of Constitutional Social Rights on Social Spending*, 60 J. L. & Econ. 713 (2017); Adam Chilton & Mila Versteeg, *Courts' Limited Ability to Protect Constitutional Rights*, 85 U. Chicago L. Rev. 293 (2018).

[97] Chilton & Versteeg, *Do Constitutional Rights Make a Difference?*, *supra* note 96.

[98] *Id.* at 582–3.

[99] *Id.* at 583.

[100] Chilton & Versteeg, *The Failure of Constitutional Torture Prohibitions*, *supra* note 96.

[101] Chilton & Versteeg, *Rights Without Resources: The Impact of Constitutional Social Rights on Social Spending*, *supra* note 96.

[102] Chilton & Versteeg, *Courts' Limited Ability to Protect Constitutional Rights*, *supra* note 96.

[103] Chilton & Versteeg, *Do Constitutional Rights Make a Difference?*, *supra* note 96, at 580 (referring to the paper of Goderis and Versteeg on the diffusion of constitutional rights, Goderis & Versteeg, *supra* note 70, at 4).

[104] *See* Section 4.1(a) above.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

legal system.[105] Therefore, Austria is coded as not having an explicit prohibition of torture even though Article 3 ECHR provides exactly that for the Austrian legal system. More subtly, rights may be also guaranteed by settled institutional practices which are not reflected in the constitutional text.[106] For example, the German Basic Law does not contain an explicit prohibition of torture either in its constitutional document.[107] However, it is undisputed that the guarantee of human dignity, which is expressly mentioned in Article 1 of the German Basic Law, also implies a prohibition of torture.[108] Therefore, these two countries are contained in the dataset as not having a constitutional prohibition of torture even though in practice they have one. These are only two examples of legal systems that we know well. It seems obvious that the more wrong classifications we find in the dataset, the more biased the results are.

Measuring the dependent variable—compliance with individual rights—is equally challenging. The authors rely on a dataset developed by Cingranelli and Richards, who base their index on a coding of the *United States Country Reports on Human Rights Practices*.[109] This approach to quantify human rights violations is fraught with fundamental conceptual difficulties.[110] It already starts with the definition of what constitutes a rights violation.[111] Courts in different states may reasonably disagree about whether a particular action violates a specific right. For example, what constitutes a violation of the freedom of expression differs greatly between the United States and many European jurisdictions. Equally, in the case of torture, debates about what practices should constitute torture are often led on the definitional level.[112]

[105] On the coding practice, see Law & Versteeg, *supra* note 72, at 1189, and the table in Goderis & Versteeg, *supra* note 70, at 9. On the constitutional status of the convention guarantees in Austria, see Christoph Bezemek, Grundrechte in der Rechtsprechung der Höchstgerichte ¶ 1.5 (2016).

[106] For the case of the United States, see, e.g., Samuel Issacharoff & Trevor W. Morrison, *Constitution by Convention*, 108 Cal. L. Rev. 1913 (2020) (calling these practices constitutional conventions).

[107] Curiously, data on the prohibition of torture and the right to establish political parties in the German constitution is not included in the study even though the number of constitutional rights is displayed, see Goderis & Versteeg, *supra* note 70, at 8–9. It seems implausible to be able to count the latter without having information on the former. Nevertheless, the point we want to make is more general in nature; the German Constitution only serves as an example.

[108] *See* Niels Petersen, Deutsches und Europäisches Verfassungsrecht II: Grundrechte und Grundfreiheiten ¶ 4.16 (2019).

[109] David L. Cingranelli & David L. Richards, *The Cingranelli and Richards (Ciri) Human Rights Data Project*, 32 Hum. Rts. Q. 401, 406 (2010). For a detailed analysis of the value of these country reports as an independent source of countries' human rights records, see Judith Eleanor Innes, *Human Rights Reporting as a Policy Tool: An Examination of the State Department Country Reports*, *in* Human Rights and Statistics: Getting the Record Straight 235 (Thomas B. Jabine & Richard P. Claude eds., 1992).

[110] *See* Robert Justin Goldstein, *The Limitations of Using Quantitative Data in Studying Human Rights Abuses*, *in* Human Rights and Statistics: Getting the Record Straight, *supra* note 109, at 35.

[111] *Id.* at 38–41. *See also* Madhav Khosla, *Is a Science of Comparative Constitutionalism Possible?* (Apr. 28, 2020) (unpublished manuscript) (on file with the authors) (pointing out that differences in performance may well be due to differences in interpretation).

[112] *See* Goldstein, *supra* note 109, at 39. Chilton and Versteeg assume this problem away by making the assumption that "constitutional prohibition of torture set the same standard everywhere": Chilton & Versteeg, *The Failure of Constitutional Torture Prohibitions*, *supra* note 96, at 428.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

Even more importantly, there is an inverse relationship between the quality of data on human rights compliance that is available and the actual human rights record.[113] The worse the human rights record of a country, the more incentives it has to obscure this fact.[114] This also means that comparisons between countries and even across time are virtually impossible because the quality of information is inconsistent.[115] Cingranelli and Richards try to mitigate this problem by only providing ballpark figures. They distinguish between three categories of human rights compliance—frequent violations, some violations, and no violations.[116]

While not fully addressing the mentioned conceptual difficulties, this approach creates problems for the type of studies Chilton and Versteeg seek to present. The three categories of rights violations are very broad and only capture fundamental changes. If Chilton and Versteeg, therefore, do not find an effect of constitutional guarantees on the level of protection of individual rights, it might be that a potential, but still practically significant, positive effect of including a right in the constitution was simply too small to be captured by the broad indicator. But even proponents of constitutional rights would probably agree that rights rarely have the revolutionary effect of turning an oppressive regime into a beacon of the rule of law.[117] Therefore, it is to be expected that any positive effect of including a constitutional right in a constitution is rather moderate.

b) Problems of identification

But the problems of the discussed studies go beyond the mere low quality of the data used. A study of the effectiveness of constitutional rights must accommodate potential selection effects. States valuing specific rights may—at the same time—be more likely to include them into their constitution and respect them to a greater degree. The authors try to address the possible selection effect through a propensity score matching.[118] They create pairs of states with a similar propensity score. The score consists of several control variables, such as the level of democracy, GDP per capita (logged), population size, the engagement in armed conflicts, judicial independence, and the number of international non-governmental organizations (NGOs).[119]

113 Goldstein, *supra* note 109, at 44–5.

114 Goodman & Jinks, *supra* note 94, at 175. For the case of torture, this is implicitly admitted by Chilton and Versteeg when they discuss why constitutional torture guarantees are ineffective, Chilton & Versteeg, *The Failure of Constitutional Torture Prohibitions*, *supra* note 96, at 422. However, this factor inhibits not only implementation but also the measurement of implementation. Nevertheless, the authors do not address the latter problem.

115 George A. Lopez & Michael Stohl, *Problems of Concept and Measurement in the Study of Human Rights*, in Human Rights and Statistics: Getting the Record Straight, *supra* note 109, at 216, 217.

116 David L. Cingranelli & David L. Richards, *Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights*, 43 Int'l Stud. Q. 407, 409 (1999).

117 *See* Adam Samaha, *Low Stakes and Constitutional Interpretation*, 13 U. Pa. J. Const. L. 305 (2010).

118 Chilton & Versteeg, *Do Constitutional Rights Make a Difference?*, *supra* note 96, at 580–2.

119 *Id.*, at 582.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

However, we have seen that propensity score matching has certain weaknesses when addressing an omitted-variable bias.[120] Matching relies on observable variables and can thus not exclude that the pairs of states differ in important, unobserved factors that may bias the result. We said that the propensity to value rights is probably a significant confounding factor. However, it is doubtful whether this concept is adequately captured by the control variables that the authors use to construct their propensity score. It is not unlikely that the propensity to value rights has a cultural component that is difficult to observe.[121]

In some of their studies, Chilton and Versteeg have added a constitutional ideal point in order to control for the tendency to value rights.[122] This approach is based on the following intuition: if two countries have exactly the same rights catalogue in their constitution except for the right in question, then the propensity to value rights should be comparable. While this might be true on a general level, it is doubtful for the propensity to value the specific right which is subject to the analysis. Imagine two countries that have the exact same rights catalogue, except that one state has included the freedom of expression while the other has not. Can we, in this case, safely assume that both societies hold freedom of expression in the same high regard?

This critique shows that any causal claim that the authors make about the effectiveness of constitutional rights is premature. This is further underlined by the fact that other statistical studies on the effectiveness of constitutional rights partly come to divergent results. Most notably, James Melton has argued in a recent working paper that constitutional rights matter in authoritarian political systems if there is at least a modest level of judicial independence.[123] Furthermore, Christian Davenport found that the existence of press freedom decreases censorship and political restriction,[124] and Zachary Elkins, Tom Ginsburg, and Beth Simmons observed that the incorporation of international human rights into domestic constitutions leads to a change in the actual human rights practice.[125] Finally, Mila Versteeg herself argued in a 2012 paper with Benedikt Goderis that constitutional rights paired with strong judicial review helped to mitigate rights violations by post-9/11 antiterrorism measures.[126]

To be sure, it would be wrong simply to assume the effectiveness of constitutional rights and judicial review. By second-guessing these widely held assumptions, the empirical research on the effectiveness of constitutional rights makes an important contribution. However, the results of these studies should not uncritically form the basis of policy recommendations. To conclude that constitutional rights and constitutional

[120] *See* Section 2.2(b).

[121] On the importance of culture for constitutional structure, see Khosla, *supra* note 110.

[122] Chilton & Versteeg, *The Failure of Constitutional Torture Prohibitions*, *supra* note 96, at 417, 431; Chilton & Versteeg, *Rights Without Resources: The Impact of Constitutional Social Rights on Social Spending*, *supra* note 96, at 728–9.

[123] Melton, *supra* note 95.

[124] Davenport, *supra* note 95.

[125] Elkins, Ginsburg, & Simmons, *supra* note 76, at 88–91.

[126] Benedikt Goderis & Mila Versteeg, *Human Rights Violations After 9/11 and the Role of Constitutional Constraints*, 41 J. Legal Stud. 131 (2012).

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023

review have limited effectiveness, and thus value, would not do justice to the complexity of the issue. These studies should therefore not be seen as final arbiters of a controversial discussion, but rather as an invitation to start a debate.

5. Conclusion

The preceding observations have shown that empirical research in comparative constitutional law is a very vibrant field of study. At the same time, we have seen that the field still faces many unresolved methodological challenges. What does this mean for the future of empirical constitutional law?

To be sure, the methodological challenges do not undermine the suitability of the endeavor as such. If comparative constitutional law scholars abandoned quantitative empirical projects, they would throw out the baby with the bathwater. However, we should broaden the approach to empirical constitutional law research and allow for more methodological pluralism. Currently, quantitative empirical research focuses predominantly on making causal claims by using cross-country panel data. However, many of these cross-country observational studies will run into the same difficulties that we have pointed out in this article.

There are several ways in which these challenges could be addressed. First, instead of focusing on the big, overarching questions, it would be more promising to focus on more specific research questions that can be addressed through quasi-experiments or qualitative case studies.[127] Context-sensitivity is key: studies on constitutional rights compliance in particular circumstances are probably more robust than a general analysis of constitutional rights compliance. Second, we need more descriptive, empirical studies; for example, studies that make representative (and not just anecdotal) claims about the actual practice of courts or the functioning of legal institutions. Such studies might be not as sexy as a cross-country panel data analysis testing a causal claim. However, they are informative, they pose fewer methodological challenges, and they may therefore be more meaningful.

Finally, we must take the results of empirical studies with a grain of salt. While social scientists go to great lengths to discuss the validity of their results, much of this discussion is lost in translation, when "social science insights" are uncritically invoked for justifying policy recommendations. Not every empirical study should immediately translate into such advice. Instead, we must be mindful that the robustness of the empirical results is ensured and that the results are not due to biased estimates, unidentified quirks of the data, or pure chance. Consequently, the results of this review do not call for less but for more carefully designed empirical research. While quantitative studies might appeal with the apparent objectivity of numbers, methodological pluralism is desperately needed.

[127] There might even be room for controlled laboratory experiments, see Armin Falk & James J. Heckman, *Lab Experiments Are a Major Source of Knowledge in the Social Sciences*, 326 Science 535 (2009); Engel, *supra* note 11; Chatziathanasiou, *supra* note 11.

Downloaded from https://academic.oup.com/icon/article/19/5/1810/6513585 by Administrative Headquarters - MPS user on 01 March 2023