

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS
FELLOWS' DISCUSSION PAPER SERIES

LARGE-SAMPLE, QUANTITATIVE RESEARCH DESIGNS FOR COMPARATIVE LAW?

Holger Spamann

Discussion Paper No. 32

9/2009

Harvard Law School
Cambridge, MA 02138

Contributors to this series are John M. Olin Fellows or Terence M. Considine Fellows in Law and Economics at Harvard University.

This paper can be downloaded without charge from:

The Harvard John M. Olin Fellow's Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

LARGE-SAMPLE, QUANTITATIVE RESEARCH DESIGNS FOR COMPARATIVE LAW?

Holger Spamann*

Abstract: A substantial body of comparative legal scholarship considers statements applicable to large, conceptually infinite numbers of countries. Such statements gain in credibility if they are supported by evidence from large samples of countries. Processing such vast evidence requires quantitative methods. Designing the requisite numerical measures of law is not straightforward, but an important insight from statistics suggests that this problem can be overcome by appropriate research design. While in practice considering more countries comes at the expense of less information per country, on balance large sample, quantitative research designs promise to yield interesting insights for comparative law.

(JEL: B40, K00, P50)

* Harvard Law School; hspamann@law.harvard.edu. This essay builds on remarks I offered at the 2009 AALS annual meeting on the panel on “The Doing Business Reports by the World Bank and the Legal Origins Thesis: Is Economics Replacing Comparative Law?” I thank the organizers Ralf Michaels and Ed Morrison for inviting me, and my co-panelists, Philipp Dann, Martin Gelter, Katerina Linos, Ralf Michaels, Mathias Reimann, Anna di Robilant, and Mathias Siems for very helpful comments. Financial support from a Terence M. Consideine Fellowship through the John M. Olin Center for Law, Economics and Business at Harvard Law School, and from the Program on Corporate Governance at Harvard Law School is gratefully acknowledged.

LARGE-SAMPLE, QUANTITATIVE RESEARCH DESIGNS FOR COMPARATIVE LAW?

© HOLGER SPAMANN 2009

Over the last dozen years, economists have introduced large-sample, quantitative research designs (hereafter LSQRD) into the comparative study of law.¹ The typical study collects some numerical measure of legal rules in a large number of countries (up to 129), and quantitatively analyzes the measure's correlation with economic outcomes on the one side, and with potential determinants of legal rules on the other.² These studies have been enormously successful in economics and policy-making circles:³ The main articles are among the most cited works in economics of the last decade; the indicators have been used in countless follow-up articles; and the World Bank's Doing Business unit handled and now continues much of the data collection.

We comparative lawyers have mostly ignored or criticized the economists' work.⁴ I myself have devoted considerable energy to demonstrating that some of the economists' key results were artifacts of measurement error in their data.⁵ Yet I shall argue here that, properly applied, the method of LSQRD promises to be quite useful for comparative law. Comparative lawyers can also render a service to other disciplines by injecting their expertise into studies of this kind.

To be able to cover dozens of countries, LSQRD often collect a narrower set of information from each country than typical qualitative studies, and always less than if all the researcher's energy had been directed at only one or two countries. In particular, LSQRD may focus on a more narrowly defined subset of legal rules, ignore more extra-legal aspects, and/or consider only one point in time. In and of itself, this loss of information is hardly desirable. LSQRD may more than make up for it, however, by increasing the amount of information in the cross-country dimension, and by the different analyses that this increase makes possible. Whether LSQRD are on balance informative will depend on the research question, the obtainable information, and on the ingenuity of the researcher. Usually, neither LSQRD nor the classical, qualitative approach will be

¹ Previous attempts to use quantitative methods in comparative law faded and did not leave much of an impact, see, e.g., John Henry Merryman, *Law and Development Memoirs II: SLADE*, 48 AM. J. COMP. L. 713 (2000).

² Cf., e.g., Simeon Djankov et al., *Private Credit in 129 Countries*, 84 J. FIN. ECON. 299 (2007).

³ For a recent survey of much of the literature, see Rafael La Porta et al., *The Economic Consequences of Legal Origins*, 46 J. ECON. LITERATURE 285 (2008).

⁴ See the description of the panel for which these remarks were written: "Much of this research runs directly counter to current comparative law scholarship Unsurprisingly, economists and comparatists have either ignored or criticized each other." For a careful endorsement of the economists' method similar to mine, see Mathias M. Siems, *Statistische Rechtsvergleichung*, 72 RABELS ZEITSCHRIFT 354 (2008).

⁵ Cf. Holger Spamann, *The "Antidirector Rights Index" Revisited*, REV. FIN. STUD. (forthcoming) (revisiting Rafael La Porta et al., *Law and Finance*, 106 J. POL. ECON. 1113 (1998)); and *id.*, *Legal Origins, Civil Procedure, and the Quality of Contract Enforcement*, J. INSTITUTIONAL & THEORETICAL ECON. (forthcoming) (revisiting Simeon Djankov et al., *Courts*, 118 Q.J. ECON. 453 (2003)).

unambiguously superior. Rather, they will shed light on the same issue from different angles.

Of course, LSQRD are not suitable for all types of research questions.⁶ Large samples of countries are informative only if the research question actually applies to large numbers of countries. This is not the case for many important research questions in comparative law, such as attempts to understand particular foreign legal systems.⁷ As I will illustrate in section I, however, much comparative law scholarship explicitly or implicitly studies theories about social phenomena that are supposed to hold true in a large and conceptually unlimited number of countries. While particular countries may serve as a prism through which to study such theories, there is no *a priori* reason to limit the analysis to these particular countries. On the contrary, large samples are the natural testing ground for any theory that is supposed to hold true in a large number of jurisdictions, as I discuss in section II. In section III, I go on to argue that quantitative methods are the only way to process what would otherwise be overwhelming amounts of information from large samples.

Consequently, I have little doubt that LSQRD are in principle desirable in comparative law. The real question is whether they can be implemented in practice, i.e., whether it is possible at reasonable cost to construct measures of the relevant phenomena that are sufficiently meaningful to generate convincing results. While some of the economists' early studies may not have reached that target in many comparatists' eyes,⁸ the latest, much more sophisticated work is more persuasive.⁹ Ultimately, the proof of the pudding will be in the eating for each individual study. In section IV, I review some research design features that, together with an important insight from statistics, suggest that the measurement problems will often be manageable. Briefly stated, the statistical insight is that imperfect measures can be satisfactory because in large samples, random measurement error will usually cancel out.¹⁰

The argument pertaining to measurement error is part of a broader point about the proper criterion for evaluating research designs. The proper criterion is not the maximization of descriptive accuracy and detail, but whether the method can detect interesting connections in the overabundance of available information. To be able to say something useful, research designs must disregard irrelevant or tangential information, and they may tolerate minor inaccuracies if these inaccuracies do not interfere with the

⁶ Cf. Detlev Vagts, *Comparative Company Law – The New Wave*, in Festschrift für Nicolas Druéy zum 65. Geburtstag 595, 595-96 (Rainer J. Schweizer et al. eds., 2002) (stressing that the “legal origins” literature is interested in different questions than much comparative law).

⁷ I have in mind an approach as set out in, e.g., Mitchel de S.-O.-l’E. Lasser, *The Question of Understanding*, in *Comparative Legal Studies: Traditions and Transitions* 197 (Pierre Legrand & Roderick Munday eds., 2003), and as applied in, e.g., *id.*, *Judicial Deliberations: A Comparative Analysis of Judicial Transparency and Legitimacy* (2004).

⁸ See, e.g., Vagts, *supra* note 6, reviewing La Porta et al., *supra* note 5. For a critique of key early studies based on a revision of their data, see Spamann, *supra* note 5.

⁹ Compare for example the studies cited *infra* note 15 and in section IV below.

¹⁰ I present this argument informally. Like other arguments in sections III and IV, however, it relies on an extensive statistical and econometric literature that makes these arguments mathematically precise.

main point and if correcting them is not worth the effort.¹¹ In principle, this is true even for the thickest of “thick descriptions.” LSQRD only make the point more obvious through their focus on large numbers of countries at the expense of the amount of detail per country. Whether this focus of LSQRD is informative must be judged for each individual study and depends to some extent on one’s prior conceptions of the world. Many sophisticated views of the world stress the importance of subtle details and hence lead to skepticism about the possibility of detecting interesting connections from the bird’s eye perspective of most LSQRD.¹² At the same time, the importance of subtle details should also induce skepticism about our ability to discern the relevant factors in analyzing small numbers of countries, and therefore only reinforces the importance of large samples that I discuss in section II. Moreover, while LSQRD may not reveal the subtle mechanisms underlying certain connections, they can at least verify or detect the connections’ existence.

In this essay, I am only concerned with the abstract merits and limitations of LSQRD for comparative law. There are a number of related questions that should be clearly distinguished.

First, LSQRD are more than numerical measurement. Some, like the World Bank, also use numerical measures of individual countries’ legal systems for comparisons to one another (rankings) or to some benchmark (benchmarking). While these may be done for large numbers of countries, the relevant comparisons are always pairwise, and the large sample is not essential. By contrast, as I explained above, the potential strength of LSQRD comes from their ability to extract information from large samples. I also mentioned that some simplification involved in measurement may be acceptable in LSQRD if and because it is outweighed by the additional complexity afforded by large samples, and because random measurement error cancels out in large samples. Rankings and benchmarking do not offer this countervailing benefit and hence place much higher demands on the quality of measurement.

Second, LSQRD are distinct from the “legal origins theory” with which they are usually associated in comparative law.¹³ Applications of LSQRD outside the “legal origins” mold pervade (comparative) sociology, political science, and economics,¹⁴ and

¹¹ For a similar point in relation to modeling, see Gillian Hadfield, *The Strategy of Methodology: The Virtues of Being Reductionist for Comparative Law*, 59 U. TORONTO L.J. 223 (2009).

¹² See, in the context of political science, Peter A. Hall, *Aligning Ontology and Methodology in Comparative Research*, in *COMPARATIVE HISTORICAL ANALYSIS IN THE SOCIAL SCIENCES* 173 (James Mahoney & Dietrich Rueschemeyer eds., 2003).

¹³ On the “legal origins theory,” see La Porta et al., *supra* note 3.

¹⁴ For two examples of particular interest to comparative lawyers, see JAMES R. BARTH ET AL., *RETHINKING BANK REGULATION – TILL ANGELS GOVERN* (2006); and ZACHARY ELKINS ET AL., *THE ENDURANCE OF NATIONAL CONSTITUTIONS* (forthcoming). In these fields, the usefulness of LSQRD is not disputed in principle, although particularly political scientists have been intensively debating the proper mix of LSQRD and other methods, *see, e.g.*, GARY KING ET AL., *DESIGNING SOCIAL INQUIRY – SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* (1994); David Collier et al., *Critiques, Responses, and Trade-Offs: Drawing Together the Debate*, in *RETHINKING SOCIAL INQUIRY – DIVERSE TOOLS, SHARED STANDARDS* 195, esp. 221-26 (on trade-offs) (Henry E. Brady & David Collier eds., 2004); Henry E. Brady et al., *Toward a Pluralistic Vision of Methodology*, 16 *POL. ANALYSIS* 353 (2006); ALEXANDER L. GEORGE &

some have already been published in this journal.¹⁵ In fact, LSQRD have also been used to criticize the “legal origins theory.”¹⁶ My only point about the “legal origins” debate in this paper is that we should neither disregard evidence merely because it comes in an unusual (quantitative) format, nor dismiss an entire method because individual results may have been erroneous or misinterpreted.¹⁷

Finally, LSQRD are not the only field for fruitful cooperation between economics and comparative law.¹⁸ Economic models can elucidate the likely effect of legal rules and thereby facilitate the application of the functional method in comparative law. Inversely, comparative evidence can challenge theoretically derived conclusions regarding the optimality or even necessity of certain domestic legal arrangements. Quantitative evidence from domestic data can provide important information for otherwise intractable comparative questions.¹⁹ And so on.

I. Statements about Large Numbers of Countries in Comparative Law

LSQRD are a method for investigating questions about large numbers of countries, particularly questions about abstractly defined groups of countries with a conceptually infinite amount of members. Explicitly or implicitly, important parts of comparative law ask precisely such questions.

Perhaps the most obvious example of an explicitly general question in comparative law is the long tradition of identifying characteristic traits that distinguish common and civil law countries. Other examples include: “[C]an legislation act as an important factor in the creation of market economies?”²⁰ Which legal systems establish

ANDREW BENNETT, CASE STUDIES AND THEORY DEVELOPMENT IN THE SOCIAL SCIENCES 3-6 (2005); Dvora Yanow, *In the House of “Science,” There Are Many Rooms*, in PERESTROIKA! THE RAUCOUS REBELLION IN POLITICAL SCIENCE 200 (Kristen Renwick Monroe ed., 2005).

¹⁵ E.g., Daniel Berkowitz et al., *The Transplant Effect*, 51 AM. J. COMP. L. 163 (2003); Nuno Garoupa & Tom Ginsburg, *Guarding the Guardians: Judicial Councils and Judicial Independence*, 57 AM. J. COMP. L. 103 (2009); John Armour et al., *How Do Legal Rules Evolve? Evidence from a Cross-Country Comparison of Shareholder, Creditor, and Worker Protection*, 57 AM. J. COMP. L. 579 (2009).

¹⁶ Cf. Berkowitz et al., *supra* note 15; Raghuram G. Rajan & Luigi Zingales, *The Great Reversals: The Politics of Financial Development in the Twentieth Century*, 69 J. FIN. ECON. 5 (2003); Mark J. Roe, *Legal Origins, Politics, and Modern Stock Markets*, 120 HARV. L. REV. 460 (2006); René M. Stulz & Rohan Williamson, *Culture, Openness, and Finance*, 70 J. FIN. ECON. 313 (2003). For counterarguments to these critiques, see La Porta et al., *supra* note 3.

¹⁷ For an extended argument that two key results of the “legal origins” literature were artifacts of measurement error, see Spamann, *supra* note 5. For an attempt to reconcile the economists’ remaining findings of differences between common and civil law countries with many comparative lawyers’ belief that there are nowadays few if any relevant *intrinsic* differences between common and civil law, see Holger Spamann, *Contemporary Legal Transplants – Legal Families and the Diffusion of (Corporate) Law*, 2009(6) BYU L. REV. (forthcoming, December 2009).

¹⁸ For a recent overview, see Ralf Michaels, *The Second Wave of Comparative Law and Economics?*, 59 U. TORONTO L.J. 197 (2009).

¹⁹ Cf., e.g., Sergei A. Davydenko & Julian R. Franks, *Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the U.K.*, 58 J. FIN. 565 (2008) (comparing the level of creditor protection in France, Germany, and the United Kingdom using data on borrowing terms, defaults, and recovery rates for small firms).

²⁰ Gianmaria Ajani, *By Chance and Prestige: Legal Transplants in Russia and Eastern Europe*, 43 AM. J. COMP. L. 93, 97 (1995).

themselves as intellectual role models for the world?²¹ Or, adding a normative dimension, what is the optimal model of judicial participation in plea bargaining?²²

On the surface, most comparative legal scholarship has more modest ambitions and considers only two or three countries explicitly: Why are there more tort suits in the United States than in the United Kingdom?²³ Why are German contracts shorter than American contracts?²⁴ Why did economic analysis of law succeed in the United States but not, or only much less so, in Germany?²⁵ To answer these questions, however, most comparativists invoke theories that reach beyond the particular countries under investigation, and developing such theories is often what motivates the comparative studies in the first place. For example, comparative lawyers investigating the questions just stated have argued that plaintiffs are more likely to sue if the cost-benefit calculation is more favorable;²⁶ that contracts tend to be shorter if good faith standards limit the payoff from, and hence the incentive to insert, one-sided terms hidden in complex contracts;²⁷ and that economic analysis of law tends to prosper in legal academic settings that have moved beyond legal doctrinalism and are open to utilitarian thought.²⁸

Such theories claim general applicability and are not limited to the countries that prompted the investigation.²⁹ This is not a coincidence but inherent in comparison (as opposed to understanding/analyzing each country on its own terms). For example, the logic of the comparative legal studies cited above is a but-for argument: if only country *A* but not country *B* experienced phenomenon *X*, it is because only country *A* but not country *B* had feature *Y*. An argument of this sort implies that a general relationship of the form “*Y* entails *X*” holds in both countries *A* and *B*. If this is so, however, there is no principled reason why the relationship should not also hold in countries *C*, *D*, *E*, etc. Of course, the postulated theory may only apply to certain types of countries, such as countries of a comparable level of development or form of government. And a researcher may legitimately judge that the detailed analysis of two countries will yield more insight than the less detailed analysis of a large number of countries. But the key point is that the underlying theory is not limited to two particular countries, even if the researcher chooses these two as a prism through which to investigate the theory.

²¹ Ugo Mattei, *Why the Wind Changed: Intellectual Leadership in Western Law*, 42 AM. J. COMP. L. 195 (1994) (book review).

²² Jenia Iontcheva Turner, *Judicial Participation in Plea Negotiations: A Comparative View*, 54 AM. J. COMP. L. 199 (2006).

²³ Basil S. Markesinis, *Litigation-Mania in England, Germany, and the USA: Are We So Very Different?*, 49 CAMBRIDGE L.J. 233 (1990).

²⁴ Claire A. Hill & Christopher King, *How Do German Contracts Do As Much With Fewer Words?*, 79 CHICAGO-KENT L. REV. 889 (2004).

²⁵ Kristoffel Grechenig & Martin Gelter, *The Transatlantic Divergence in Legal Thought: American Law and Economics vs. German Doctrinalism*, 31 HASTINGS INT’L & COMP. L. REV. 295 (2008)

²⁶ Markesinis, *supra* note 23, at 252 (highlighting contingency fee arrangements, no loser-pay rules, and punitive damages in the United States).

²⁷ Hill & King, *supra* note 24, at 890-91.

²⁸ Grechenig & Gelter, *supra* note 25.

²⁹ *Cf.* Lee D. Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 2-4 (2002) (arguing that much legal scholarship implicitly makes broader empirical claims).

Besides, generalizing beyond the observations under study is necessary if we want the experience of some countries to inform policy choices in another. In the example above, to harness the experience of countries *A* and *B* for the benefit of country *C* requires a theory going beyond countries *A* and *B*.³⁰ This is not an argument for “one size fits all” policy recommendations,³¹ or for pretentious theories claiming certain knowledge of the world. Theories can be about the process of policy-making rather than the content of the resulting regulation.³² Theories can explicitly or implicitly incorporate local context, yielding different policy recommendations for different settings. Theories can be limited to explaining a part of the relevant variation and hence be open to overriding considerations. And credible theories in the comparative context will be probabilistic rather than deterministic (see section III below). But unless the theories dare to generalize something, they cannot bridge the gap between the observations that form the basis of our knowledge and the new country, time period, etc., for which we want to use that knowledge.³³

II. The Case for Large Samples

When a statement explicitly or implicitly applies to large numbers of countries, large samples are the natural source of evidence to support the statement. Evidence from only one or two countries is unlikely to convince the skeptic that the evidence is representative of the larger group to which the statement applies.

This is not an abstract, philosophical concern. There is considerable danger, documented in other disciplines, that the properties of a few salient observations are erroneously taken as representative of some larger group.³⁴ Worse, researchers may consciously or unconsciously seek out precisely those observations that fit their theory.³⁵ Such selection bias almost inevitably arises if the observations under study are the very

³⁰ Even if we only wanted past experience in countries *A* and *B* to inform future policy choices in these two countries, we would need a theory going beyond the past experiences of countries *A* and *B*—and again, there would be no principled reason why a theory should link past experience of *A* to future events in *B* but not *C*.

³¹ About the dangers of such recommendations, see, e.g., Katharina Pistor, *The Standardization of Law and Its Effect on Developing Economies*, 50 AM. J. COMP. L. 97 (2002).

³² For example, a theory could assert that solutions developed locally under specified conditions are on average superior (by some specified measure) to global templates. Cf. Berkowitz et al., *supra* note 15 (arguing that the mode of reception of legal transplants is much more important than the content of the transplant).

³³ Cf. Dani Rodrik, *The New Development Economics: We Shall Experiment, But How Shall We Learn?*, in *WHAT WORKS IN DEVELOPMENT? – THINKING BIG AND THINKING SMALL* (Jessica Cohen & William Easterly eds., forthcoming) (emphasizing the need for empirical results to be generalizable to other settings, while advocating a country-specific, diagnostic approach to development policy that “does not imply absence of theory. The only meaningful way in which one can sift through the evidence—or indeed know what kind of evidence to look for—is through the prism provided by clearly articulated frames.”).

³⁴ On the detrimental effects of this phenomenon for theory development in comparative politics, see, e.g., BARBARA GEDDES, *PARADIGMS AND SAND CASTLES – THEORY BUILDING AND RESEARCH DESIGN IN COMPARATIVE POLITICS* (2003). For a description of the proper approach to the selection of observations and other issues of inference specifically in legal applications, see Epstein & King, *supra* note 29.

³⁵ See, e.g., Barbara Geddes, *How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics*, 2 POL. ANALYSIS 131 (1990).

observations that inspire the theory.³⁶ An obvious example of such an error in comparative law would be to take unique features of the contemporary United States, such as the heavy use of class actions and the civil jury, as representative of contemporary common law jurisdictions in general. We avoid such errors by considering more observations from the class of countries to which the statement purports to apply (“common law jurisdictions”).

In addition, evidence from only a handful of countries will often not be conclusive on its own terms. The fewer observations, the more theories will be able to “explain” the observed pattern. To rule out more theories, or to assess their relative importance, requires more observations (countries) exhibiting different combinations of possible explanatory factors.

III. The Necessity of Quantitative Methods

The analysis of information from large samples almost inevitably requires quantitative methods. Perhaps the best evidence for this is that large collections of comparative legal information, such as Abel’s and Lewis’s impressive collection on the role of lawyers in society, have hardly been utilized in classical comparative law.³⁷ There are two reasons why qualitative methods alone cannot fully exploit the richness of such collections.

The first reason is that most relevant hypotheses will be probabilistic (“tend to,” “more likely,” etc.), and hence their verification requires the calculation of averages from the data—a quantitative operation, albeit a very basic one.³⁸ By contrast, if the relevant hypotheses were deterministic, we would only need to check if they are contradicted by some individual observation, and we might be able to do that by a “manual” review of the available information.³⁹ Unfortunately, we do not understand the world well enough to formulate credible deterministic answers to the type of questions I discussed in section I. Probabilistic knowledge is usually the most we can hope to achieve. Take, for example, the question investigated by the economists’ first application of LSQRD in comparative law: does legal investor protection, particularly on-the-books investor protection, foster financial market development?⁴⁰ Financial market development in the United Kingdom up to the middle of the twentieth century arguably shows that such protection is not

³⁶ See, e.g., KING ET AL., *supra* note 14, at 20-21.

³⁷ By August 22, 2009, LAWYERS IN SOCIETY (3 vols., Richard L. Abel & Philip S.C. Lewis eds., 1988-89) had been cited only two times in the American Journal of Comparative Law (Lexis search for “Abel w/s Lewis w/s “lawyers in society””). Also cf. Merryman, *supra* note 1 (lamenting that the wealth of data assembled in the SLADE project was never analyzed), and Vagts, *supra* note 6, at 596 (noting that the Encyclopedia of Comparative Law’s volume on Business Associations has received little attention).

³⁸ It is possible, however, that a particular observation seems so strongly determined by a probabilistic theory (i.e., the observation is the theory’s ideal scenario) that a failure of the theory’s prediction for this observation would considerably reduce our confidence in the theory, see Timothy McKeown, *Case Studies and the Limits of the Quantitative Worldview*, in RETHINKING SOCIAL INQUIRY, *supra* note 14, 139, 158-62.

³⁹ Cf. Ronald Rogowski, *How Inference in the Social (but Not the Physical) Sciences Neglects Theoretical Anomaly*, in RETHINKING SOCIAL INQUIRY, *supra* note 14, 75, 77-79 (reporting how case studies of one observation successfully challenged clearly articulated theories).

⁴⁰ La Porta et al., *supra* note 5.

necessary,⁴¹ and failure of such development in Russia after adoption of new investor protection rules in the mid-1990s arguably shows that it is not *sufficient*.⁴² And yet, few would conclude from these observations that legal investor protection on-the-books is irrelevant for financial market development. It may still be the case that financial markets *on average* prosper more when investor protection on the books is better, and this might be useful information for a policy maker deciding how many resources to allocate to improving such rules. Provided we can measure such investor protection (see the following section IV), we could then start the analysis by calculating whether countries with good investor protection on average have more developed financial markets.

This example also illustrates the second, even more important reason for the use of quantitative methods: simultaneously to keep in play for dozens of countries various relevant factors that come in shades of grey rather than only black and white. In the example above, we would not only want to know whether on-the-books investor protection is on average helpful for financial markets, but by how much. After all, there are not just protective and unprotective rules, or developed and undeveloped markets, but varying degrees of protection and development. Moreover, there are many other factors that presumably have a bearing on financial market development, such as a country's general level of development, its distribution of wealth, or its accounting rules, and we want to distinguish the effects of these factors from those of on-the-books investor protection. To detect even simple relationships between these various non-binary factors in dozens of countries is more than the human brain can handle unassisted. To detect and assess patterns in such data—to distill the useful from the excess of information available—we need quantitative methods.

Quantitative methods do not perform magic. The principal tool for the analysis of comparative data, linear regression, estimates partial *correlations*, and provides measures of the precision of the estimates (partial correlations are correlations between two variables after removing the part that is correlated with, and hence potentially accounted for by one or more other variables⁴³). While regression and other statistical techniques can establish *causation* under certain assumptions,⁴⁴ these assumptions will rarely if ever be convincing for comparative data.⁴⁵ For this and other reasons, the correlations

⁴¹ See, e.g., Brian R. Cheffins, *Does Law Matter? The Separation of Ownership and Control in the United Kingdom*, 30 J. LEGAL STUD. 459 (2001).

⁴² See, e.g., Bernard Black et al., *Russian Privatization and Corporate Governance: What Went Wrong?*, 52 STAN. L. REV. 1731, esp. 1753 (2000) (noting that Russia did adopt good corporate and bankruptcy law in 1996 and 1998, respectively).

⁴³ For an introduction to linear regression, see, e.g., CHRISTOPHER DOUGHERTY, INTRODUCTION TO ECONOMETRICS chs. 1-9 (2d ed. 2002).

⁴⁴ See for the philosophical basis Paul W. Holland, *Statistics and Causal Inference*, 81 J. AM. STAT. ASS'N 945 (1986); and for a recent review of the available strategies JOSHUA D. ANGRIST & JÖRN-STEFFEN PISCHKE, MOSTLY HARMLESS ECONOMETRICS: AN EMPIRICIST'S COMPANION (2009).

⁴⁵ See, e.g., Jonathan Klick, *The Perils of Empirical Work on Institutions*, J. INSTITUTIONAL & THEORETICAL ECON. (forthcoming); Philip A. Schrod, *Beyond the Linear Frequentist Orthodoxy*, 16 POL. ANALYSIS 335 (2006). On the trade-off between interesting questions and rigorous demonstration of causation generally, see, e.g., Rodrik, *supra* note 33.

estimated by LSQRD need to be supplemented with other theoretical and empirical insights, including, of course, those of qualitative studies.

IV. Measuring Law?

The discussion so far has presupposed that the requisite data to operationalize quantitative methods in comparative law exist, or at least can be created in the future. In reality, the possibility of creating sensible numerical measures of legal concepts is highly controversial. Attempts to measure law in the “legal origins” literature and the World Bank’s Doing Business project are considered a major innovation by some and a major foolishness by others.⁴⁶ The critics might agree that “when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind,”⁴⁷ but point out that it is better to have such “meagre and unsatisfactory” knowledge than to have no knowledge at all disguised in meaningless numbers.

The complexity of law in general, and the difficulties of ascertaining foreign law in particular, certainly suggest that creating comparative measurements of law is not easy. But other disciplines, particularly the social sciences, operate with similarly complex concepts (e.g., culture; welfare policies), and yet the usefulness of quantitative methods is not disputed in those fields (even though there is considerable disagreement regarding the appropriate balance between quantitative and qualitative work).⁴⁸ If the experience of those other disciplines is any guide, we almost surely stand to gain by increasing the use of quantitative methods in comparative law. There are several ways of dealing with the measurement problem:

First, depending on the research question, it may be possible to use an indirect quantitative assessment of legal institutions and thereby to circumvent direct measurement of law. For example, to study the importance of public securities law enforcement for stock market development, it is possible to look at the budget of public securities regulators rather than the rules governing the regulators’ powers and composition.⁴⁹ Similarly, instead of attempting to gauge minority investor protection from on-the-book legal rules and perhaps information about their enforcement, we may be able to use some financial indicator that we believe to be a direct function of minority investor protection, such as the discount on minority shares compared to control block trades.⁵⁰

⁴⁶ For a critical review at least of the actual implementation of the approach, see, e.g., Bertrand du Marais, *Les limites méthodologiques des rapports Doing Business*, in *DES INDICATEURS POUR MESURER LE DROIT?* 17, 35-64 (Bertrand du Marais ed., 2006).

⁴⁷ WILLIAM THOMSON [Lord Kelvin], 1 *POPULAR LECTURES AND ADDRESSES* 80 (2d ed. 1891).

⁴⁸ See, e.g., the references cited *supra* note 14. On the trade-offs involving data quality in particular, see, e.g., Collier et al., *supra* note 14, at 204-09, 225.

⁴⁹ See Howell E. Jackson & Mark J. Roe, *Public and Private Enforcement of Securities Laws: Resource-Based Evidence*, 93 *J. FIN. ECON.* 207 (2009).

⁵⁰ For such a measure, see Alexander Dyck & Luigi Zingales, *Private Benefits of Control: An International Comparison*, 59 *J. FIN.* 537 (2004).

Second, even direct measurement of law need not be complicated. For example, comparative lawyers have been interested in the effects of contingency fees.⁵¹ Whether they are allowed or not and, if so, under which conditions, would be relatively easy to determine and to encode into simple, possibly binary variables.⁵²

Third, more complex legal arrangements can often be summarized by one or more pertinent intermediate outcomes. For example, the quality of bankruptcy law can be measured by recording bankruptcy duration and recovery rates for a well-defined set of facts.⁵³ Such data are comparable across countries and able to account for all interactions of legal rules within the data's domain. In particular, they integrate substantive law and enforcement. This approach is similar to the common core (or Trento) project approach in comparative law, in that it also focuses on the respective legal systems' ultimate treatment of particular fact patterns rather than on abstract rules. At the same time, the common core studies tend to explore more varied fact patterns and more basic outcomes, suggesting that comparative lawyers could have much to add to quantitative comparative legal studies in the future.

Crucially, statistics teaches us that numerical measures need not be perfect to obtain unbiased results, as long as the imperfections are not correlated with a variable of interest.⁵⁴ The simple intuition is that random errors usually cancel out in large samples. Standardized test scores, such as SAT or GRE scores, can illustrate the point. For any given student, the score will be influenced not only by the ability that the test attempts to measure, but also by physical well-being on the day of the test, ease with taking standardized tests, and sheer luck. Moreover, it may be the case that the concept of ultimate interest, say "intelligence," is richer than the ability measured by the test, and that even though the measured and unmeasured components tend to go together, one or the other may be relatively more pronounced for some individual. Thus we should hesitate to compare two students based only on their respective test scores. We need to hesitate much less, however, to compare two *groups* based on their *average* scores, and the larger the groups, the less we need to hesitate. After all, both groups are likely to include students that are ill, uneasy, or unlucky, or whose "intelligence" is not well represented by the tested ability, and the larger the groups, the more likely they do so in equal proportions. This evaluation would change only if we had reason to believe that disturbing factors did not affect the groups equally, for example if fire alarms went off repeatedly the night before the test in the dormitory of only one of the groups.

⁵¹ See, e.g., Markesinis, *supra* note 23.

⁵² For a simple indicator of the permissibility of contingency fees, see Djankov et al., *supra* note 5.

⁵³ See Simeon Djankov et al., *Debt Enforcement Around the World*, 116 J. POL. ECON. 1105 (2008); THE WORLD BANK, *DOING BUSINESS 2009* (2008).

⁵⁴ Technically, this statement only holds for measurement error in the dependent variable. Measurement error in the independent variable biases the regression estimate towards zero, i.e., will make the correlation between two variables appear less strong than it really is (for an illustrative example, see the main text below accompanying note 55); in multivariate settings, the effects are more complicated. Econometricians have, however, devised techniques for dealing with these issues. See, e.g., JEFFREY M. WOOLDRIDGE, *ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA* 63-76, 105-07 (2002).

Consequently, LSQRD can often content themselves with rough numerical measures that would be much too coarse for the comparison of only two countries. A particularly important corollary is that recording the treatment of a few standardized fact patterns may be sufficient to measure an entire area of law for purposes of LSQRD. All that is required is that the treatment of these particular fact patterns is indicative of countries' remaining law in the area, and that there is no group of countries identifiable by the other variables of interest for which the scenarios are particularly unrepresentative. For example, depending on the research question, it may be possible to derive a satisfactory measurement of the speed and cost of civil procedure in general from data on the speed and cost of civil procedure in only a few typical cases, such as a simple contract dispute and a simple tort case (e.g., a traffic accident). This considerably reduces the effort required to measure law for LSQRD.

I do not claim that the quality of measurement is irrelevant for LSQRD. First, while random errors would surely cancel out in an infinite sample, they will not fully cancel out in a finite sample—and the smaller the sample and the larger the individual errors, the greater the remaining aggregate error is likely to be. In the test score example above, the two groups are probably roughly equally affected by random illness, unease, or bad luck, but this is not certain, especially if the groups are small. Second, random measurement error tends to make the relationship between two variables appear weaker than it really is, even if it does not erroneously suggest either a negative or a positive relationship.⁵⁵ For example, if one explored the relationship between temperature and the likelihood of water freezing with temperature measurements that are randomly 10 degrees too high or too low, one would still find that freezing is more likely in low temperatures, but one would erroneously find some positive probability of freezing above 0°C and of not freezing below 0°C.

Reliable measurement requires considerable care and resources. To derive its Doing Business indicators, the World Bank employs dozens of dedicated staff and collects information from thousands of lawyers around the world.⁵⁶ This seems necessary because the key results of less elaborate early studies turned out to be artifacts of measurement error.⁵⁷ While individual researchers can hardly match the World Bank's resources, networks of comparative lawyers, such as those assembled for the common core project, could collect equally reliable or even more reliable data.

V. Conclusion

In this short essay, I have argued that we comparative lawyers should try to harness the power of LSQRD for the benefit of comparative law, and to contribute our expertise to improve quantitative comparative legal studies in other fields. Some may wonder whether LSQRD using legal input still belong to the field of comparative law. In my view, this is of secondary importance. What matters is whether such studies can help answer relevant questions, and whether they benefit from the input of researchers with

⁵⁵ See WOOLDRIDGE, *supra* note 54.

⁵⁶ Cf. <http://doingbusiness.org/MethodologySurveys/MethodologyNote.aspx>.

⁵⁷ See Spamann, *supra* note 5.

knowledge of the law of different countries and with an appreciation for the difficulties of comparison. I think the answer is yes on both counts.

I do not believe LSQRD should or could “replace comparative law,” as suggested by the title of the panel for which these remarks were written. As I said at the beginning, there are major branches of comparative law that are not amenable to LSQRD, and even within the branches that are, LSQRD must be sensibly combined with qualitative methods. I do believe, however, that we comparative lawyers stand to gain from incorporating LSQRD into our arsenal of methods.