Unpacking the Rule of Law: A Review of Judicial Independence Measures[*]

Julio Ríos-Figueroa
Assistant Professor
División de Estudios Políticos, CIDE
Carretera México-Toluca 3655
Col. Lomas de Santa Fe, 01210, México D.F.
Tel: (52 55) 57279800 x. 2128
julio.rios@cide.edu

Jeffrey K. Staton
Assistant Professor
Department of Political Science
Emory University
1555 Dickey Drive
Atlanta, Georgia 30322
W: (404) 727-6559
F: (404) 727-4586
jeffrey.staton@emory.edu

March 12, 2009

Abstract

The rule of law has become a favored solution to crucial substantive problems in the social sciences. Several measures have been created to capture the concept; however, we lack a good sense of which of the numerous ones are appropriate to particular research projects. We argue that a broad, multidimensional concept of rule of law should be abandoned and replaced with the individual concepts from which it is constituted. Focusing on judicial independence, a central component of the rule of law concept, we evaluate the validity of thirteen measures that have been used in the literature. We find considerable support for the validity of a number of *de facto* measures of independence, less support for *de jure* measures, and strong evidence that *de facto* and *de jure* measures are tapping into distinct concepts. Finally, we highlight a significant missing data problem in these data and suggest how this problem is compounded by common practices used to demonstrate robustness.

[*]Prepared for Caltech-USC Center for the Study of Law and Politics and University of Texas Law School Symposium, *The Rule of Law*, March 26-27, 2009. Previous versions of this paper were presented at the 2008 European Consortium of Political Research (ECPR) Joint Sessions Workshop, at the 2008 American Political Science Association (APSA), and at the 2009 Workshop on Law, Politics and Human Rights (Emory University).

Over the last quarter century, scholars have come to view the rule of law as a solution to crucial substantive problems in the social sciences. Elements of the rule of law are afforded great causal weight in theoretical models of economic development and growth (Barro 1997), state capacity (North and Weingast 1989; Stasavage 2002), the protection of human rights (Apodaca 2004; Cross 1999; Keith 2002) and even political and social order (North and Weingast 1989; North, Summerhill and Weingast 2000). Building the rule of law has become a goal of non-governmental organizations, major state powers and international institutions of all types (e.g. Carothers 2006). It is perhaps little exaggeration to say that the rule of law has become a sort of mantra for those interested in promoting human welfare via institutional design. Reflecting its conceptual importance, the number of rule of law measures has increased tremendously in recent years. Haggard et al (2008) report on seven. We review thirteen in this paper. Clearly, we do not want for data. In principle, as we accumulate more and more information, we find ourselves in a better position to test theoretical claims about the rule of law and to track legal reform. In practice, however, we lack a good sense of which of the numerous available measures are appropriate in particular research projects. Which of these measures reflect well the theoretical concept of interest?

Unfortunately, scholars have expressed serious concerns about measurement validity. Haggard et al (2008) note that existing measures are not strongly related to each other (some negatively so), and this is true even when we evaluate only measures that were designed to tap into the same conceptual feature of the rule of law (e.g. security of contract). In addition, there is also a significant problem of data availability. Whereas the Political Risk Service's (PRS) measure is available for a long time series, the World Bank's measure only dates back to the mid-1990s. Many other measures are available cross-sectionally only (e.g. Feld and Voigt 2003; La Porta et al 2004). Scholars also have raised concerns about the varied, potentially unreliable, and likely biased processes by which the measures are generated. Most directly, subjective measures of the rule of law (be they derived from surveys of firms, entrepreneurs, experts, or citizens) have prompted some critics to ask whether the measures are biased towards the interests of the surveyed: "Because international experts often tended to be businessmen, theirs was a one-

sided view and measure. Singapore, for example, ranks at the top of many ratings" (Hammergren 2006, 14).

Although these features of the measurement landscape are easily demonstrated and suggest reasonable concerns, they obscure the central problem of evaluating the validity of rule of law measures, which is simple, and frankly, well-known (e.g. Haggard et al 2008). Scholars do not share a definition of the rule of law. It is an "essentially contested concept" (Waldron 2002). Of course, the contest is reflected in the measures. With its aggregate indicator, the World Bank seeks to capture "the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence" (World Bank 2007, 3). The Fraser Institute attempts to measure the "protection of persons and their rightfully acquired property," and the PRS produces a single index that supposedly reflects the strength and impartiality of the legal system and popular observance of law.[1] Without a shared concept, it is not clear how to compare validity across measures. So what is the rule of law then? Does it mean societal trust in legal institutions? Popular compliance with formal legal rules? Judging that is unaffected by extra-legal concerns? Good contract enforcement? Does it imply all of these things? If so, do some of them matter more than others, and if so, in what ways exactly? In so far as measurement validity claims are always tied intimately to the concepts the indicators supposedly reflect, this problem could not be any more pressing.

Scholars are aware of this problem; however, we typically proceed without addressing it directly. We either note conceptual differences across measures and then evaluate measurement validity, the differences notwithstanding (e.g. Skanning 2008); or, we note the differences and then suggest that further research is required into the processes by which distinct strands of the rule of law are related to one another (e.g. Haggard et al 2008). In this paper, we adopt a different approach. We wish to offer three simple, practical suggestions about how to pick and subsequently analyze rule of law measures.

[1] See Gwartney and Lawson (2007) for the Fraser Institute's measure. The PRS's measure description at: http://www.prsgroup.com/ICRG_Methodology.aspx.

First, at the risk of taking a position on concepts, which we generally view as unassailable, the rule of law concept in its broadest sense is simply too big for the vast majority of research projects. As we discuss below, the broad concept contains at least three dimensions: 1) an institutional dimension that captures the extent to which government can be constrained by other elements of the state, 2) an individual dimension that captures the extent to which individuals are treated fairly by the legal system, and 3) a social dimension that captures political order and social stability. And what's more, each dimension contains a number of distinct subcomponents (e.g. judicial or bureaucratic independence, access to courts, procedural due process, etc.). There are many reasons to caution against such a broad concept, but we wish to stress that it is unhelpful for two reasons. It makes it difficult to learn precisely what about the rule of law matters and what does not. The concept includes distinct subcomponents that common social scientific theories view as causally related (e.g. government constraint is a cause of social order and judicial impartiality is a cause of judicial trust). These are theoretical claims, which require testing but which are true by definition under a broad rule of law concept. Perhaps more importantly, existing conceptual definitions do not clearly articulate how to aggregate information across the three dimensions, much less across subcomponents within one dimension. Unless we are willing to assume that all dimensions and subcomponents should be weighted equally, it is near impossible to evaluate the validity of any rule of law measure. For the same reason that an unclear theory cannot be falsified, an unclear concept cannot be validated. Still, it is reasonable and highly useful to develop stand-alone measures of the components captured under the rule of law umbrella. In this paper, we will focus on a subcomponent of the first dimension. We will seek to evaluate the validity of measures of *judicial independence*, a central component of the first dimension, and one that plays an important theoretical role in numerous literatures. Far from the gloomy picture suggested by the lack of tight correlations across measures described by Haggard et al, we suggest cautious optimism for judicial independence measures.[2]

[2] This optimism does not suggest, however, that estimating the effects of judicial independence from existing measures is a trivial enterprise, especially in light of the obvious endogeneity concerns Haggard et al raise (also see Acemoglu, Johnson, and Robinson 2001).

Second, we must let our theoretical models, within which our concepts operate, guide our measurement choices. Collier and Adcock (1999, 539) argue that "specific methodological choices are often best understood and justified in light of the theoretical framework, analytic goals, and context of research involved in any particular study." We could not agree more. Scholars are no doubt familiar with common distinctions between *de jure* and *de facto* judicial independence, but there are important differences between different concepts of *de facto* independence, as well; and these differences matter greatly for what you wish to measure. What is more, we have good theory and evidence that judicial behavior is strategic, and for that reason, we might like a measure that is robust to the observable effects of judicial strategy (e.g. Epstien and Knight 1998; Helmke 2005; Murphy 1964; Vanberg 2005). We hope to suggest a few ways by which this might be done.

The third implication derives from the nature of data availability. Although it is clear that the states and years for which data are available vary wildly across measures, what is less well understood is that the missing data problem is broader and more consequential than that measures simply return no information on some states in some years. Even within a block of years for which measures produce values, missingness patterns are non-random, which implies an important challenge to inference that empirical researchers really should not ignore. As we discuss below, this pattern not only raises concerns about bias in particular models, it also suggests that common strategies for demonstrating robustness are likely to reinforce those biases. We present simple evidence of non-random missingness for twelve of the thirteen measures we evaluate below, and we have no reason to believe that our thirteenth measure is necessarily immune to the problem. Fortunately, there are well-known statistical corrections for the inferential biases that emerge when we fail to take account of non-random missingness (Horton and Kleinman 2007). For this reason, scholars who use these data ought to be particularly concerned about the process by which they deal with missing data.[3] We divide the remainder of the paper into two substantive

[3] The construct validity tests we present in this draft may suffer from the biases we note here. We are currently working on the imputation problem, which has been difficult to crack in two ways. First, there are two general processes by which missing data emerge. One of them deals with why some measurement teams could not obtain information on a state-year for which they intended a score. The other deals with why some teams elected to target

sections and a conclusion. We first address the broad rule of law concept and then turn to an evaluation of judicial independence. We conclude with some implications of our main points.

II. The Rule of Law Concept

We lack a standard definition of the rule of law. Common definitions usually include a list of elements, longer or shorter depending on the purposes of the writer, which is said to be necessary for the rule of law. Thus, for instance, Joseph Raz argues that "the rule of law requires the law to be prospective, open, clear, and relatively stable. That the passing of laws be governed by open, permanent, clear, and general rules. And, that the independence of courts is guaranteed" (Raz 1977, 198. See also Macedo 1994, 148). Even though there is a list of requisites, it is not clear how the elements in the list relate to each other, which ones may be necessary conditions for others, or if a subset of the elements (or a certain amount of all of them) is sufficient to consider that law rules in a given country.

While there is no consensus on what the rule of law is, there seems to be agreement on what the rule of law stands against. As Judith Shklar put it, three negative elements can lead to a better understanding of an otherwise obscure concept of the "rule of law": "the insecurity of arbitrary government, the discriminations of injustice, and the fear of violence" (Shklar 1998, 37). These elements suggest three dimensions that are part of the rule of law: an institutional dimension linked to the contention of governmental arbitrariness, an individual dimension linked to the prevention or correction of discriminatory practices in law enforcement, and a social dimension related to instability and violence.

For many theorists, the institutional dimension is of crucial importance. As Macedo (1994, 149) writes, "[M]ore than anything else, the rule of law stands against arbitrariness and caprice: it regularizes political power and renders it impersonal." It is also within this dimension where we have a strong tradition of theoretical and practical arguments that have been used in the design of institutions. In classic institutional theory, dividing the state into concurrent and mutually dependent powers has been thought

---

some states but not others, or some years but not others. It is not yet clear to us how to model these different sources of missingness simultaneously. The second problem is that, even if we assume that there is one underlying model of missingness, we have been unable to obtain convergence using any of the imputation algorithms in Amelia (King, Honaker, Joseph, and Scheve 2001). It is certainly possible that the missingness is significant that we will not be able to obtain multiple imputation estimates.

crucial to precluding arbitrariness and promoting liberty (Montesquieu 1997; Locke 1998; Madison et al 1961). Within the institutional dimension, the judicial component has been considered paramount historically (e.g. Montesquieu 1787, Book II); and, modern institutional research suggests that independent judiciaries that constrain arbitrary state power ensure that state promises to respect individual rights are perceived credible (e.g. North 1990). This credibility breeds efficient investment and growth (e.g. Frye 2004). Researchers also have suggested that well-functioning legal systems are critical to the protection of human rights (Cross 1998; Keith 2002).

An institutional structure that prevents arbitrariness, however, does not exhaust the concept of rule of law. The individual dimension relative to discrimination in law enforcement is also of central importance. As Montesquieu warned us: "it is not sufficient to [examine] political liberty as relative to the constitution; we must examine it likewise in the relation it bears to the subject", because "the constitution may happen to be free, and the subject not" and vice versa (1997, Book XII, 216). And the "political liberty of the subject … consists in security, or in the opinion people have of their security", and this security "is never more dangerously attacked than in public or private accusations" (1997, Book XII, 217). Here Montesquieu was thinking about criminal laws, the due process of law, and even the customs that determine what we can more generally call the *ideal of equality before the law*. This ideal includes not only security concerns but also equality of treatment regarding positive rights.

Finally, prominent theorists such as Judith Shklar have also considered a social dimension related to instability and "the fear of violence" as part of the rule of law concept. This third dimension covers features related to whether order or open conflict, violence, upheavals, and crime are present in a country and can undermine the rule of law. At the extreme the pervasiveness of political violence that threatens the very existence of a state calls for the temporal suspension of the rule of law, which according to some hampers the state's reaction to secure its survival (see Finn 1991). Also the systematic presence of riots, kidnappings, political terror, or organized crime generally produce calls for detentions without trial, searches without warrants, torture, and other measures that violate central features of the rule of law.

Thus, in its most general sense, the rule of law involves three dimensions, within which fall distinct and varied components. Table 1 summarizes this structure.

[Table 1]

Given the breadth of this concept, it is not surprising that attempts to measure it throw as wide a net as possible. The most obvious example of this approach is the World Bank's Rule of Law measure, which aggregates survey questions that tap into components as distinct as access to land and access to water for agriculture, the pervasiveness of money laundering through banks, confidence in the police and the neutrality of the judges, and violent and organized crime (Kaufmann, Kraay, and Mastruzzi 2007b, 74). Importantly, there is no clear theoretical statement of the relative weight we should place on the separate dimensions and components. Is criminal due process twice as important as judicial trust? Half as important? Does social order matter at all if government is entirely arbitrary? Since there is no clear theory suggesting how to aggregate this information, purely atheoretical procedures are the norm.[4] Such approaches do not take into account potential relations between dimensions of the rule of law, let alone potential relations between components within those dimensions. For this reason alone, it is near impossible to evaluate the validity of a rule of law measure. How can we decide whether a measure is properly aggregating information if we do not know conceptually how we are supposed to aggregate?

In another sense, the situation is analogous to the substantive/procedural debate in the context of democracy studies. Like substantive concepts of democracy, the broad rule of law concept has come to mean all normatively good aspects of political life, aspects of life we would like to explain and which other dimensions (and components) of the concept could advance. The substantive conception of democracy is credited with features such as representation, accountability, equality, participation, dignity, among others. As Przeworski and his coauthors argue, "From an analytical point of view, lumping all good things together is of little use. The typical research problem is to examine the relationships among

[4] The World Bank is based on the use of an unobserved components model on the data from different sources. The values of WB are weighted averages of the underlying data, with weights reflecting the precision of the individual data sources. See Skanning (2008) for a discussion of aggregation procedures or other measures.

them" (2000, 14). A minimalist or procedural conception of democracy allows the researcher to empirically assess whether core elements of democracy process actually produce equality, development or human satisfaction. Substantive conceptions simply assert that such relationships exist (Dahl 1971; Przeworski et al 2000, 14).

We believe that the same problem infects the rule of law. We might like to advance social order and we have theory suggesting that good legal institutions designed to constrain government can help do so (North, Summerhill and Weingast 2000). How can we know if this theory is wrong if we insist on lumping together distinct pieces of a causal process into one measure? We believe that abandoning the multidimensional rule of law concept and focusing on its dimensions or on components within each dimension will lead to more fruitful research. The big questions over whether the rule of law generates economic development, human rights protection, or corruption control can be more fruitfully explored if broken down into the dimensions of rule of law. Does judicial independence protect human rights or advance development? Does social trust in the impartiality of judicial officers reduce corruption? Under what conditions does protecting property rights produce higher levels of investment? These are the questions we should be asking. We should use measures designed to let us answer them.

This line of argument suggests a simple implication for our efforts to evaluate the validity of rule of law measures. We ought to limit the exercise to particular dimensions, and better still, to components within those dimensions. By doing so, we give ourselves the best chance to link a measure to a coherent and empirically useful concept. In the remainder of this paper, we focus on the concept of judicial independence, a critical component of the first rule of law dimension.

III. Judicial Independence Measures

Conceptual Matters

Like the literature on the rule of law, conceptual differences characterize the literature on judicial independence (Burbank and Friedman 2002; Russell and O'Brien 2001)[5]. Still, we can distinguish

[5] There may be subtle but important differences regarding independence concepts depending on to whom independence refers (for example, the highest court, a court, a judge, the entire judiciary) and independence from

between two general ways in which scholars use the term. The first concept of independence demands that judges be the "authors of their own opinions" (Kornhauser 2002, 42-55). On this account, a judge is independent when she does not face undue external or internal pressure (as say from hierarchical superiors) to resolve cases in particular ways.[6] In other words, a judge is independent when the output of judicial processes reflects sincere judicial preferences (Becker 1970, 1-8). Judicial independence in this sense is *judicial autonomy*. What judges think is what they produce.

A second concept of judicial independence takes into account a fundamental problem of judicial policy making. As the 78th Federalist reminds us, lacking financial or physical means of coercion, courts depend on the assistance of other political authorities to enforce their decisions (Madison et al 1961). Under this second concept, the argument is that it makes little sense to call a judge independent if her decisions are routinely ignored or poorly implemented. Judicial independence requires not only that judges resolve cases in ways that reflect their sincere preferences, but also that their decisions are enforced in practice (Cameron 2002; Larkins 1996). Judicial independence in this sense means *judicial power*. What judges think is what they produce and what they produce controls the outcomes of legal conflicts. In so far as the first dimension of the rule of law is about constraining government, it seems likely that the second concept, which implies state constraint, is largely what scholars have in mind. Indeed, the United Nations principles of judicial independence require that judicial decisions are respected (United Nations 1985). That said, it is clearly possible for a particular theory to invoke a judicial autonomy concept and we will demonstrate below that measures of judicial independence attempt to measure autonomy.

The distinction between judicial autonomy and judicial power operates on what scholars have previously called a *de facto* concept of judicial independence. A more familiar distinction in the literature

---

whom (for example, other political branches, superior in the judicial hierarchy, parts in a trial). In this paper, however, we focus on the broader conceptual ideas.

[6] Of course, political authorities always apply some pressure when they are parties to a case or when they file briefs as interested parties (e.g. amicus curiae briefs and other similar institutions). The key here is that this pressure is perfectly acceptable within the rules of the legal system.

involves the difference between *de facto* and *de jure* concepts (e.g. Feld and Voigt 2003; Rios-Figueroa 2006). The latter deals with formal rules that are designed to insulate judges from undue pressure, either from outside the judiciary or from within. In our view, it is probably misleading to distinguish, at the conceptual level at least, between *de facto* and *de jure* "judicial independence." Judicial independence is best thought of as either autonomy or power, whereas institutions like fixed tenure, budgetary autonomy, and judicial councils are best thought of as rules designed to promote autonomy or power. They do not, strictly speaking, reflect judicial independence any more than, say, single-member district, plurality voting rules reflect disproportionate political representation or any more than multiple veto players reflect political stability. It is certainly possible that rules of judicial insulation enhance judicial independence, but these are theoretical relationships between distinct concepts that need to be clearly established and tested. It seems conceptually preferable to call these rules what they are: institutions of judicial insulation. That said, since the literature persists in calling rules of insulation *de jure* judicial independence, we will do so the same here.

Up to this point, we have addressed three types of judicial independence concepts. Before proceeding, however, we wish to consider how our measurement strategies might address the behavioral implications of strategic models of judicial decision-making. This is useful, we think, whether one adopts an autonomy or power concept of judicial independence. Simmons (2007, 22), writing about the advantage of litigation strategies to advance human rights norms, provides a representative of the kind of behavior we have in mind. She writes:

> One of the most important conditions for litigation to be a potentially useful strategy to enforce rights is judicial independence. For courts to play an important enforcement role, they must be at least somewhat independent from political control. The government or one of its agencies, representatives or allies is likely to be the defendant in rights cases, and unless local courts have the necessary insulation from politics, *they are unlikely to agree to hear and even less likely to rule against their political benefactors*. Anticipating futility, individuals or groups may decide to avoid the courts altogether.

The kind of strategic judicial deference alluded to here is anticipated by common models of judicial-government interaction, and has received considerable empirical support in studies around the globe (Ginsburg 2003; Helmke 2005; Herron and Randazzo 2003). Strikingly, scholars have even uncovered strategic effects on courts generally thought to be highly independent, like the United States Supreme Court and the Federal Constitutional Court of Germany (Martin 2006; Vanberg 2005). If these models are correct, then non-systematic observation of judicial decision-making, litigation strategies and compliance as a means of developing measures of independence can be significantly misleading. Indeed, a court that offers little constraint on government can appear to be highly constraining if it chooses its cases wisely. It certainly can ensure systematic compliance, which makes inferring judicial power from the outcomes of legal conflicts difficult. For this reason, it will be useful to consider how our measures might deal with strategic judicial behavior. We return to this issue below.

Content Validity Analysis

We have identified thirteen measures that social scientists have used in empirical research. Table 2 partitions the measures into categories of *de jure* and *de facto* independence and indicates the years for which they are available. Further, within the category of *de facto* independence, we have indicated (whenever it is clear) whether the researchers responsible for the measure intended to capture an autonomy or power concept. Our appendix contains detailed information on the construction of the measures summarized in Table 2. As is clear from the table, far more measures aim at *de facto* independence than *de jure* independence, and within the former category, scholars largely seem to have targeted the power concept.

[Table 2]

A systematic account of content validity would require a thorough evaluation of the extent to which each measure reflects the theoretical concept of interest. Ultimately, this is a subjective exercise, as there is always something to quarrel with when considering the content of particular measures, especially so when concepts are contested and subject to strategic behavior. Rather than provide a systematic account of the content validity of each measure, we wish to provide some representative examples,

highlight a few curious choices, and finally suggest how measures might address the strategic behavior problem. The appendix provides the information necessary for a reader to evaluate content independently.

The Howard and Carey (2004) measure provides a representative example of a *de facto* independence measure aimed at the concept of judicial autonomy. Howard and Carey (2004, 286) define judicial independence as, "The extent to which a court may adjudicate free from institutional controls, incentives, and impediments imposed or intimidated by force, money, or extralegal, corrupt methods by individuals or institutions outside the judiciary, whether within or outside government." Their measure, which is derived from United States State Department country reports, is a three-category ordinal measure, which indicates whether the high court of a state in a particular year is fully independent, partially independent or dependent. According to Howard and Carey (2004, 287-88), a state has a fully independent judiciary if the high court functions in practice:

- [Independently] of the executive and legislature, and is
- Relatively free from corruption and bribery and provides basic criminal due process protections to criminal defendants.

A state has a partially independent judiciary if its high court either satisfies the first or the second condition or partially satisfies both. A state has a dependent judiciary if its high court satisfies neither condition.

In general, the measure is reasonably well connected to the concept. That said, Tate and Keith (2007) question whether Howard and Carey's focus on criminal due process rules moves the measure away from standard concepts of judicial independence. If Howard and Carey have the autonomy concept in mind, then this concern is certainly apt. It is quite possible for a court to be the "author of its own opinions" even if it provides little due process protections for criminal defendants. Indeed, the court under analysis could be fully responsible for eliminating such protections through its doctrinal choices. And it is certainly possible that in some states, the legal infrastructure necessary to advance the kind of actions necessary to liberalize criminal due process doctrine does not exist. If this is true, then no matter how independent a court is, and no matter how much it might like to extend the rights of the criminally

accused, it simply will not be confronted with the cases useful to making such a rights revolution (Epp 1998). In such cases, the Howard and Carey measure will be biased toward dependence.

Henisz (2000) provides a representative example of a power, *de facto* independent measure. Henisz wishes to measure the extent to which a state's judiciary serves as a constraint on government. To that end, the Henisz measure incorporates information from two other indicators of the power, *de facto* independence concept: the Polity Executive Constraints (Polity) measure and the Political Risk Service's Law and Order (PRS) measure. The Polity measure indicates the extent to which a state's chief executive faces constraints on its decision-making process. According to the Polity authors, one (but not the only) constraint on this authority is an independent judiciary. The PRS measure combines a law component, intended to measure the "strength and impartiality of the legal system," with an order component, intended to measure popular observance of formal legal rules. The Henisz indicator assigns a score of 1 to states that are coded sufficiently high on both the Polity and PRS measures, and a 0 otherwise (see appendix for coding rules). The key content validity concern here is that the Henisz measure includes elements of alternative dimensions of the rule of law. Popular observance of law is not well connected to the institutional constraint dimension. That said, the logic of the Henisz measure is that on average, combing the ICRG and Polity measures should reflect *de facto* independence better than either the Polity or ICRG measure does on their own. There is no evidence that this is true.

The chief *de jure* measures include Feld and Voigt's survey-based score and Keith's (2002) index, used by Apodaca (2004). Feld and Voigt produce an interval level measure of independence, derived from surveys of country experts. Their measure aggregates information on twelve formal rules that govern the constitutional status of a state's highest court, its appointment and retention rules and its judicial review authority (see appendix). Unfortunately, the measure is available only for the year 2003 and for a limited set of states, and it is based not on the authors coding of legal rules but on experts surveyed about those legal rules in their respective countries. Keith (2002) produces seven ordinal measures based on the United Nations principles of judicial independence. The scores indicate whether a state's constitution formally guarantees the tenure of high court judges; ensures the finality of judicial

decisions; grants judges exclusive authority over their jurisdiction; bans special or military jurisdiction for civilian crimes; financial autonomy; a separation of powers system; and specifically enumerates appointment qualifications. For each component, the Keith measure returns a value of 2 if a constitution explicitly provides a feature of *de jure* independence; 1 if it does so partially or vaguely; and, 0 otherwise. In the case of military jurisdiction, Keith also includes a code of -1 in the event that a constitution explicitly endorses military jurisdiction. Apodaca (2004) creates an additive index of these seven items, whose range is from -1 to 14. Tate and Keith (2007, 23) have questioned the validity of such a scale, noting that at least two factors reflect the underlying dimension over which these measures vary.

Hybrid Measures and Partial Judicial Independence

Although we found evidence that the *de facto* measures seem reasonably well targeted at *de facto* concepts, it is not uncommon for these measures to include institutional components, as well. Howard and Carey (2004) and Tate and Keith (2007), both of which are derived from the State Department reports combine behavioral and formal institutional information to produce their *de facto* measures.[7] The most curious example is Cingranelli and Richard's (2008) measure, again derived from the State Department reports. Like Howard and Carey and Tate and Keith, the Cingranelli and Richards measure provides a three category ordinal independence score (0=none; 1=partial; 2=full). To be coded as fully independent, a state's judiciary must satisfy the following criteria:

- It has the right to rule on the constitutionality of legislative acts and executive decrees.
- Judges at the highest level of courts have a minimum of seven-year tenure.
- The President or Minister of Justice cannot directly appoint or remove judges and the removal of judges is restricted (e.g. allowed for criminal misconduct).
- Actions of the executive and legislative branch can be challenged in the courts.
- All court hearings are public.
- Judgeships are held by professionals.

[7] In particular, Tate and Keith incorporate information on whether a state includes a constitutional provision for *habeas corpus* (or an equivalent legal action) and public trials, *de jure* features that the authors demonstrate to be associated with behavioral outcomes that reflect *de facto* independence.

Notice that these criteria *all provide de jure information*. Although the distinction between partial and no judicial independence seems to be based on *de facto* criteria, Cingranelli and Richards reserve the highest category for states that satisfy this set of formal rules. It is not clear why. In addition, even though *de jure* indicators are included there is no consideration of the theoretical relations between them. For instance, appointment procedures and length of tenure may be related. If the president directly appoints judges but they hold their positions for life, should they be considered less independent than judges that are appointed through non-political means but have short tenure periods?

Judicial Strategy

Once we recognize that judicial decision-making is strategic (e.g. Helmke 2005; Vanberg 2005), the core measurement challenge is that direct observation of court behavior without estimating how judicial decision-making is influenced by political concerns can be misleading. It is possible to observe an ultimately ineffective court whose decisions appear to be free from undue influence and always implemented, precisely because courts can strategically select and decide cases to minimize conflict (e.g. Ginsburg 2003). For this reason, a precise measure of effectiveness derived from court behavior requires a wealth of case specific data, which allows for the systematic estimation of the extent to which judicial decisions respond to external, political pressures (e.g. Helmke 2005; Vanberg 2005) and the extent to which judicial decisions are properly implemented (e.g. Spriggs 1996). For a worldwide sample, comparative judicial politics has yet to produce such data.[8] In light of this problem, scholars might look to measures of effectiveness that capture the types of behaviors we should observe if the judicial system functions as a genuine constraint on the state, especially behaviors that are not obviously correlated with the dependent variable. This is not to say that court specific measures are invalid. Rather, the point is that given central findings in judicial politics on strategic judicial behavior, we might consider being creative about our measurement choices until we can systematically estimate judicial effectiveness around the world.

[8] Tate et al. (2002) and Martin et al (2008) have begun to develop a broadly comparative data set on high court decisions; however, at present it covers only eleven states. Thus we are a long way away from having worldwide data on judicial systems.

One possible solution involves seeking a measure of the behavioral consequences of judicial independence, which are not the subject of analysis. For example, Powell and Staton (2009) have used the *Contract Intensive Money* (CIM) measure created by Clauge et al. (1999). The *CIM* is "the ratio of non-currency money to the total money supply" (Clague et al. 1999, 188). Conceptually, high values of *CIM* reflect a society's trust in judicial institutions that enforce the banking industry's contractual obligations. To be sure, the *CIM* was conceptualized as a measure of legal protections for property rights, and this is how it has been traditionally used in the literature (e.g. Clague et al. 1999, 186; Souva, Smith, and Rowan 2008). In order to use it as a general measure of *de facto* independence, we must assume that, on average, states that possess judicial institutions that protect property rights are likely to have judicial institutions that protect rights generally. There are good reasons to believe that *CIM* is a valid measure of the extant to which courts protect rights generally, including human rights. In particular, simple predictive validity tests indicate that the *CIM* is negatively associated with a variety of state human rights abuses, including extrajudicial killings, political imprisonment, and disappearances and it is correlated with other measures of the concept (Powell and Staton 2009). It is even associated with decisions of a state to adopt without reservation the international *Convention Against Torture* (Powell and Staton 2009). Importantly, this measure should not be sensitive to the kind of strategic judicial decision making discussed above, because it does not directly reflect case decisions or compliance, but rather picks up the behavior we should observe from individuals if courts constitute constraints on the state. On the other hand, it is obvious that the *CIM* can certainly be conceptualized as a direct measure of confidence in banking, and so if a research question requires that we distinguish between, say, confidence in banking rules and courts, then obviously, the *CIM* will be problematic. Nevertheless, even in such a situation, the challenge will again be to address the problem of strategic behavior in whatever measure is adopted. In our view, a pragmatic approach is useful here. The *CIM* or some other measure based on the behavioral consequences of a *de facto* independent judiciary, may be extremely helpful in a particular research context – in others, it may simply not make sense.

Convergent Validity Analysis

Convergent validation considers whether different measures of similar concepts are related to one another (Adcock and Collier 2001, 540). Ideally, we would like a benchmark measure of judicial independence against which we can simultaneously compare our measures. Of course, in the judicial independence literature there really is nothing close to a benchmark. That said, we can evaluate the extent to which measures are associated with each other. A quick glance at four states raises some flags.

Table 3 shows four *de facto* and two *de jure* judicial independence scores across four states, normalized to the unit interval, for the year 2003. Note that for Denmark and Germany, the measures report similar information – both countries appear to have highly independent judiciaries. But now consider Mexico. Whereas the PRS and Henisz return extremely low scores, the Fraser measure produces a middling value, and the Feld and Voigt measure is relatively high. Similarly, the *de jure* measures tell us drastically different information about Mexico in 2003. Now consider Pakistan. Here, the Fraser and Henisz scores are low, whereas the PRS returns a middling value. These results are troubling, and we will return to them below.

[Table 3]

Table 4 presents a correlation matrix for the thirteen judicial independence measures. In general, we show Pearson's correlation coefficients, but where two ordinal measures are being compared, we indicate Kendall's *Tau b.* We are most concerned with the relationships within the categories of *de facto* and *de jure* independence, as we do not necessarily expect a high correlation between *de jure* and *de facto* measures. The table suggests some simple, not necessarily unexpected, yet relatively encouraging results. The correlations between the *de facto* measures are all positive but one, and some are quite large. As we might expect, the State Department-based measures are highly correlated, but even measures derived from entirely distinct sources demonstrate a strong positive relationship (e.g. Fraser and Tate and Keith). Of course, as Haggard *et al* (2008) suggest, there is considerable variation in the degree of association, and some measures seem only very weakly correlated. For example, the correlation between Polity and Feld and Voigt (de facto) measures is only .2 and there is essentially no relationship between Feld and

Voigt (de facto) and La Porta et al – .02. Moreover, the La Porta et al and the CIM measures are actually negatively correlated – -.23. That said, these are not typical results, as the table makes clear. In general, the measures are positively correlated. Of greater concern is the evidence for the convergent validity of the *de jure* measures where for instance the Keith and Feld and Voigt (de jure) are quite weakly correlated –.18.[9]

[Table 4]

Reflecting Feld and Voigt's (2003) own analysis, the relationship between the *de jure* and *de facto* measures is tenuous at best. From a measurement validity perspective, this is actually a positive result, as the measures should be tapping into distinct concepts. From a theoretical perspective this result may reinforce the point that formal institutions of judicial insulation are not necessarily related to actual judicial independence (see Herron and Randazzo 2003), or raise a more interesting question about the conditions under which formal institutions of judicial insulation impact the autonomy of judges or the enforcement of their decisions. We might not necessarily expect a direct relationship between *de jure* and *de facto* independence, but rather expect that institutions of judicial insulation are important within particular political contexts. For example, it is possible that insulation matters more when courts cannot be easily ignored or decisions poorly implemented (Martin et al 2007), and that the institutions of judicial insulation that are placed in the constitution have a stronger impact than those placed in the organic laws of the judiciary depending on the rigidity of the constitution and the partisan composition of the organs necessary to make constitutional amendments (Pozas-Loyo and Ríos-Figueroa 2006). Moreover, it is even possible that formal institutions simply do not create unambiguous political incentives. Indeed, Helmke and Staton (2009) show how increasing judicial tenure can produce competing incentives for insincere judicial behavior, such that without knowing precisely how a judge weighs these incentives, it is possible

[9] The correlation analysis is based on all available data, from all years in which the measures report scores. In some cases, it is not possible to produce a correlation coefficient, because there is no overlap between the measures. For example, the Feld and Voigt measures are available only for the year 2003 whereas the Apodaca scale ends in 1996. To match the construct validity analysis that follows, we took the 1996 scores for our measures that were available in that year and matched those observations with the observations from 2003 for the measures available in that year (i.e. Feld and Voigt, La Porta et al, Fraser).

for greater tenure to advance more or less judicial independence.

Construct Validity Analysis

Construct validation involves considering whether measures of a concept are related to measures of alternative concepts in ways that are anticipated by theoretical models (Adcock and Collier 2001, 537). In this section, we consider two types of construct validity tests.

Human Rights

Scholars have found considerable evidence linking judicial independence to a state's protection of human rights – the physical integrity of the person in particular (Cross 1999; Blasi and Cingranelli 1996; Keith 2002; Powell and Staton 2009). In this section, we replicate Apodaca's study of human rights, using each of our thirteen measures. Apodaca herself posits that *de jure* independence should be positively associated with human rights outcomes in so far as institutions of judicial insulation create the incentives necessary for courts to constrain governments from violating fundamental human rights protections. Powell and Staton (2009) have found evidence linking *de facto* judicial independence to human rights outcomes, state sponsored torture in particular. Although the precise theoretical arguments linking judicial independence to human rights vary across scholars, there is general theoretical consensus that judicial independence should be negatively associated with poor human rights behavior – certainly it should not be positively correlated.

Apodaca's dependent variable is an average of two Political Terror Scales (Gibney, Cornett and Wood 2006), one calculated from United States State Department reports and the other from Amnesty International country reports. Each scale produces a five-category ordinal ranking of human rights abuses in a state. Apodaca takes the average score across the two scales, entering the value of one scale if the other scale is missing. She then regresses the natural log of this average on measures of a state's judicial independence, population, population density, democracy, armed conflict, economic growth and military expenditures. Believing that there is a lagged relationship between her independent variables and her outcome variable, Apodaca regresses the 1998 average PTS scale on 1996 scores for the independent

variables.[10] Replicating this strategy is straightforward where data is available in 1996. Where this was not possible because our measures are only available in 2003 (e.g. Feld and Voigt and La Porta *et al*), we regressed the 2005 PTS average on the 2003 values of the independent variables.

Table 5 summarizes our findings by the types of measures, indicating the OLS estimate for each measure, its White/Huber standard error, an indication of statistical significance in a two-tailed hypothesis test, and the sample size. Full results are available upon request, but reflect well the findings Apodaca describes in her paper.[11] All of the *de facto* measures produce negative coefficients and all those available in 1996 are statistically significant. Interestingly, however, the partially independent category for the State Department-based measures is not significant in any model. Models in which we collapse the measure into a dichotomous score (0 if none; 1 if partial or full) produce negative and significant estimates. The Keith *de jure* scale, which Apodaca used, did not reach statistical significance, and the coefficient is actually positive. Importantly, Apodaca notes in her paper that she excluded observations on "developed" states, though we do not know how "developed" is defined operationally. We did recover the negative and significant Apodaca result when restricting our sample to states with GDP/cap below US$1000. For the measures not available in 1996, we again find support with the Fraser judicial independence measure. That said, neither the Feld and Voigt nor the La Porta *et al* measures produce significant results. And like the Keith scale, the Feld and Voigt *de jure* coefficient is actually positive. On balance, the *de facto* independence results are highly encouraging, whereas the *de jure* results are not. That said, like the convergent validity tests above, it is important to recognize that from a validity perspective, it is actually encouraging that the *de facto* and *de jure* measures perform differently. The models may not suggest much support for a *de jure* independence effect on human rights outcomes; however, they certainly

[10] We do not wish to quarrel with Apodaca's estimation strategy or model specification. Our models are identical to Apodaca's in all but two ways. First, Apodaca also includes the World Bank's general rule of law index, which we exclude. The rule of law variable clearly includes judicial independence and so we estimate this effect twice by including the rule of law. Second, we include a contemporaneous measure of armed conflict. While it is plausible there are lagged effects of growth, the effect of conflict on human rights abuses seems highly likely to be contemporaneous.

[11] Because the publication is in *Judicature*, we do not have access to the actual estimates, but rather just a verbal description.

provide no reason to believe that the *de jure* measures are invalidly measuring their concepts. Rather, this finding is useful to question what precise institutional feature may be related to human rights protection that the de facto measures are capturing but the de jure measures are not. For instance, human rights protection may be more closely related to institutions of judicial insulation other than tenure of high or administrative court judges (such as in La Porta et al's measure) but to the extent of judicial review powers of judges or the existence of an independent public prosecutor's office. De jure measures tend to focus on a handful of institutions of judicial insulation, and it may be the case that measures based on an extended and careful coding of a more complete set of de jure institutions correlates better with human rights outcomes.

[Table 5]

European Settler Mortality Rates

We provide one last construct validity test, leveraging an instrument for "good institutions," developed by Acemoglu, Johnson and Robinson (2001, AJR). AJR set out to address the plausible endogeneity concern in rule of law models of economic development. It is hard to believe that quality legal institutions cause economic development but are not impacted themselves by development. Causality almost certainly runs in both directions. Consequently, a typical model of development that includes a legal institutional variable on the right hand side is likely to produce a biased estimate of institutional effects, in part capturing the effect of development on institutions. AJR instrument for these institutions with a measure of colonial settler mortality rates in the 19th century. The logic behind the instrument is that where Europeans faced environments that were relatively disease-free and generally hospitable, they built institutions that reflected those at home anticipating significant migration. These institutions were designed to protect rights. However, where they faced relatively inhospitable environments, European states did not construct such institutions as they did not anticipate significant migration. These initial choices locked in differences in institutional quality over time.

Table 6 summarizes the results of twelve regression models. We regressed our judicial independence measures on the log of the AJR settler mortality rate. We estimated ordered logit models

for the three State Department-based measures, a logit model for the Henisz measure and OLS for the remaining measures. As the table indicates, all of the *de facto* independence coefficients are negatively signed and statistically significant, an important result given the small sample size. Reflecting the general pattern of results so far, although each of the *de jure* measures is negatively signed, only the Feld and Voigt score is statistically significant. As before, the results are stronger for the *de facto* measures.

[Table 6]

Patterns of Missingness

Table 2 underscores a well-understood feature of the judicial independence data – measures are missing data for much of the 20th century. Only the Henisz, Polity, PRS and CIM measures offer anything close to a reasonable time series. For this reason, we are largely left with time invariant measures. In so far as theories of judicial independence commonly predict significant temporal change or at least change across political institutional arrangements that vary with time (e.g. the number of veto players), cross-sectional measures are simply not adequate. This much, of course, is understood. A more interesting, less well-publicized issue concerns missingness patterns in these data within blocks of years for which measures are available. For example, the BTI project excludes all OECD states, so that it almost certainly produces a judicial independence measure that is missing systematically. Indeed, the BTI should produce a non-random pattern of missingness by design. If there is significant evidence of non-random missingness across our measures, and scholars do not address it (and we have no evidence that researchers do), then it is likely that *all* estimates of judicial independence effects are subject to the biases associated with various haphazard strategies of imputation (King *et al* 2001). How serious is this problem for the measures generally? It turns out that it is quite serious.

There are many possible patterns of non-random missingness. We test about the simplest theory of missingness that we could construct, suggesting that missingness can be explained by a state's level of economic development. We anticipate that the average GDP per capita of states with missing data for judicial independence should be lower than the average where the data is not missing. Table 7 shows difference of means tests for economic development, as measured by GDP/capita in constant US dollars

in the year 2000 and its natural log. As is clear, there is evidence of non-random missingness for every judicial independence measure except the Howard and Carey measure and the differences are substantial in some cases. For the Fraser measure, average income is $9856 where the data is not missing yet plummets to $2433 where the data is missing. Moreover, although nearly every measure exhibits a missingness pattern where we are more likely to observe a score for wealthier states, the Take & Keith measure resembles the BTI score in that it is more likely to be missing for rich states.

[Table 7]

In one sense, it is perhaps not critical to note this non-random pattern of missing data. This is a feature of all kinds of data, and scholars generally have a responsibility to deal with it as it emerges. That said, it is important to stress that scholars in this field *are not dealing with it*, so that it is unclear what to make of our empirical inferences. Further, if you pair the missing data problem with a problem that emerged above in the convergent validity analysis, we have a very serious problem on our hands.

Recall from Table 3 that our selected measures seemed to perform similarly for Germany and Denmark but not for Mexico and Pakistan. Is it possible that the measures are more likely to agree for better developed states than underdeveloped states? Table 8 suggests that they are. This table shows correlation coefficients between the Take & Keith measure and all of the other *de facto* measures used in the construct validity tests, divided by whether states fall above or below the median GDP/capita level for our entire dataset (US$1550). That pattern is striking. For every measure, the correlation is considerably stronger for above the median than it is below. Indeed, for the Feld & Voigt measure, the correlation is actually negative below the median and positive above! Providing this information for all measures is cumbersome, but the results generalize. The *de facto* measures are more strongly related among developed states than they are among underdeveloped states.

Does this matter? Consider two empirical facts. First, judicial independence data is missing systematically by level of development. Second, the measures themselves are move likely to agree among the move developed states. One central recommendation of Haggard *et al* is to run robustness checks on models that use rule of law data. While we agree with this suggestion in principle, if we do this without

dealing with the missing data problem, we bias ourselves toward confirming our inferences. The reason is that we will be more likely to be conducting our robustness checks on more developed states, as these are the states for which we are more likely to have data; and, these are also the states where out measures are more likely to agree. Applied researchers simply must address the missing data problem in this context.

Summary

First, although the various judicial independence measures do not share content perfectly, there is evidence that they are all aiming at one of two well-known independence concepts – autonomy and power. We also find reasonable evidence in a variety of contexts for the convergent and construct validity of a wide array of *de facto* judicial independence measures. These measures seem to be picking up reasonably an underlying concept of independence. That said, these measures are not always strongly correlated and conceptually it is not clear that they all address strategic behavior. We provided one example of a measure that can address this problem, the CIM. For this reason, a reasonable conclusion is that since there is decent evidence for the validity of these measures generally, scholars should choose the indicators that satisfy their theoretical goals.

Second, we find that *de jure* and *de facto* measures of judicial independence do not covary. If we think that they are tapping into different concepts, the low correlations are a positive result from a measurement validity perspective. On the other hand, if we have theoretical reasons to believe that they are related, this finding calls for a refinement of two issues of research. In one sense, there is a measurement issue. Perhaps we require *de jure* measures that are better targeted at the broad set of incentives that judges face rather than a few institutional rules like tenure and appointment. But we think the result calls for better theoretical models, which identify more precisely which and how institutions for judicial insulation produce judicial independence.

Finally, we stress that scholars must seriously consider the non-random patterns of missingness in their data, especially in this context. Robustness analysis, which should provide us all with greater confidence in the arguments we produce, is no simple fix here. Indeed, simply substituting one judicial independence measure for another and considering whether results "hold up" is biased toward finding that

they do.[12]

Conclusion

The concept of rule of law is now regularly found in arguments about its importance to life in peaceful, free, and prosperous societies. Yet the lack of a standard definition should make us cautious about the several measures that measure it. We argued that the rule of law concept in its broadest sense is simply too big for the vast majority of research projects, and that a broad definition impedes the theoretical and empirical assessment of causal relations between its different dimensions and components. Focusing on judicial independence, a crucial component or the rule of law, is a step ahead in using narrower concepts but, still, how do we choose amongst different measures of judicial independence? We argued that theory should be the guide in making this choice. Distinguishing between three concepts of judicial independence we evaluate the validity of thirteen measures that have been used in the literature and find reasonable evidence in a variety of contexts for the convergent and construct validity of a wide array of *de facto* judicial independence measures. In contrast, *de jure* measures do not converge and we think this finding poses interesting questions for future versions of the paper. For instance, should we aim at coding the broad set of incentives that judges face rather than a few institutional rules like tenure and appointment? Under what conditions de jure institutions for judicial insulation have an effect on de facto measures of judicial behavior? Or is it that the case that formal institutions of judicial insulation are not necessarily related to actual judicial independence?

We also find interesting non-random patterns of missing data across all measures. Although it is clear that data availability varies wildly across measures, what is less well understood is that the missing data problem is broader and more consequential than that measures simply return no information on some states in some years. Missingness is related to levels of development and this coupled with convergence of de facto measures of judicial independence call into question the adequacy of robustness analysis. We

[12] Of course, we would be remiss if we did not address the obvious tension in our conclusions. We have not addressed it ourselves in this working paper, our strategy for dealing with missingness in our validity tests (i.e. listwise deletion) is far from adequate, and raises questions about the uncertainty we need to have about our validity analysis. See discussion in footnote 3 above. We intend to address this problem in a subsequent version. At this point, we simply note that we should approach the results with an appropriate degree of skepticism.

of course recognize that the non-random missingness, and the challenge to inference that it implies, also means that the validity analysis carried out in this paper should be taken with a grain of salt. In future versions of the project we plan to derive multiple imputation estimates of the parameters and re-do the validity tests.

References

Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *The American Political Science Review* 95 (3): 529-546.

Apodaca, Clair. 2004. "The Rule of Law and Human Rights." *Judicature* 87 (6): 292-299.

Bertelsmann Transformation Index. 2008. *Manual for Country Assessments.* Munich, Germany: BTI. http://www.bertelsmann-transformation-index.de

Becker, Theodore. 1970. *Comparative Judicial Politics.* Landham, MD: University Press of America.

Burbank, Stephen B. and Barry Friedman (eds.). 2002. *Judicial Independence at the Crossroads. An Interdisciplinary Approach*, California: Sage.

Cameron, Charles M. 2002. "Judicial Independence: How Can You Tell It When You See it? And, Who Cares?" In *Judicial Independence at the Crossroads: An Interdisciplinary Approach*, ed. Steven B. Burbank and Barry Friedman. New York: Sage Publications Inc, 134–147.

Carothers, Thomas. 2003. "Promoting the Rule of Law Abroad. The Problem of Knowledge". *Carnegie Endowment for International Peace*, Working Paper # 34.

Cingranelli, David L. and David L. Richards. 2008. *The Cingranelli-Richards (CIRI) Human Rights Data Project Coding Manual Version 2008.3.13.*

Clague, Christopher, Philip Keefer, Stephen Knack, and Mancur Olson. 1999. "Contract-Intensive Money: Contract Enforcement, Property Rights, and Economic Performance." *Journal of Economic Growth* 4 (2): 185-211.

Collier, David and Robert Adcock. 1999. "Democracy and Dichotomies. A Pragmatic Approach to Choices about Concepts". *Annual Review of Political Science*, no. 2, pp. 537-565.

Cross, Frank B. 1999. "The Relevance of Law in Human Rights Protection." *International Review of Law and Economics* 19 (1): 87-98.

Davis, Kevin. 2004. "What Can the Rule of Law Variable Tell us About Rule of Law Reforms?," *Michigan Journal of International Law*, vol. 26, no. 141.

Epstein, Lee and Jack Knight. 1998. *The Choices Justices Make*. Washington D.C.: CQ Press.

Feld, Lars P., and Stefan Voigt. 2003. "Economic Growth and Judicial Independence: Cross Country Evidence Using a New Set of Indicators", *European Journal of Political Economy*, vol. 19, 497-527.

Finn, John E. 1991. *Constitutions in Crisis: Political Violence and the Rule of Law*. New York: Oxford University Press.

Gibney, Mark, Linda Cornett and Reed Wood. "Political Terror Scale 1976-2006." August 19, 2008, from the Political Terror Scale Web site: http://www.politicalterrorscale.org/

Ginsburg, Tom. 2003. *Judicial Review in New Democracies: Constitutional Courts in Asian Cases.* New

York: Cambridge University Press.

Gurr, Ted R. *Polity Ii: Political Structures And Regime Change*, 1800-1986 [Computer file]. Boulder, CO: Center for Comparative Politics [producer], 1989. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1990.

Gwartney, James, and Robert Lawson. 2007. *Economic Freedom of the World. Annual Report*. The Fraser Institute. Available at: http://www.freetheworld.com/2007/EFW2007BOOK2.pdf

Haggard, Stephan, Andrew MacIntyre and Lydia Tiede. 2008. "The Rule of Law and Economic Development." *Annual Review of Political Science 11: 205-234.*

Hammergren, Lynn. 2006. *Envisioning Reform*. Philadelphia: Penn State University Press.

Helmke, Gretchen. 2005. *Courts under Constraints: Judges, Generals, and Presidents in Argentina*. Cambridge: Cambridge University Press.

Henisz, Wittold. 2000. "The Institutional Environment for Economic Growth", *Economics and Politics*, vol. 12, no. 1, March, pp. 1-31.

Helmke, Gretchen and Jeffrey K. Staton 2009. "Courting Conflict: The Logic of Risky Judicial Decisions in Latin America." Working Paper. URL: http://www.juecesypoliticaenamericalatina.cide.edu/

Herron, Erik S. and Kirk A. Randazzo. 2003. "The Relationship Between Independence and Judicial Review in Post-Communist Courts." *Journal of Politics* 65 (May):422-438.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2007. "Governance Matters VI: Governance Indicators for 1996-2006". *World Bank Policy Research Working Paper* No. 4280

Keith, Linda Camp. 2002. "Constitutional Provisions for Individual Human Rights (1977-1996): Are They More than Mere "Window Dressing?" *Political Research Quarterly* 55 (1): 111-143.

King, Gary and James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49-69.

Kornhauser, Lewis A. 2002. "Is Judicial Independence a Useful Concept?" In *Judicial Independence at the Crossroads: An Interdisciplinary Approach*, ed. Steven B. Burbank and Barry Friedman. New York: Sage Publications Inc, 134–147.

Larkins, Christopher M. 1996. "Judicial Independence and Democratization: A Theoretical and Conceptual Analysis." *American Journal of Comparative Law* 44 (4): 605-626.

La Porta, Rafael, Florencio Lopez-de-Silanes, Christian Pop-Elches, and Andrei Shleifer. 2004. "Judicial Checks and Balances", *Journal of Political Economy*, 112, no. 2.

Locke, John. 1998. *Two Treatises of Government*. New York: Cambridge University Press.

Macedo, Stephen. 1994. The Rule of Law, Justice, and the Politics of Moderation. In *The Rule of Law*, ed. Ian Shapiro. Nomos XXXVI. New York: New York University Press.

Madison, James, Alexander Hamilton, and John J. Jay, J. 1961. *The federalist papers*. New York: New American Library.

Martin, Andrew D. 2006. "Statutory battles and constitutional wars : Congress and the Supreme Court." In *Instituitonal Games and the U.S. Supreme Court*, James R. Rogers, Jon R. Bond and Roy B. Flemming, eds. Charlottesville: University of Virginia Press.

Martin, Andrew D., Clifford J. Carrubba, Matthew Gabel, Gretchen Helmke, and Jeffrey K. Staton. Collaborative Research: A Cross-national Study of Judicial Institutionalization and Influence. National Science Foundation Grant. Law and Social Sciences Program (SES-0751340)

Montesquieu. 1977. *The Spirit of the Laws* (*A Compendium Of The First English Edition. Edited With An Introduction By David Wallace Carrithers*). California: University of California Press.

Munck, Gerardo y Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices". *Comparative Political Studies*. Vol. 35, No. 1, pp. 5-34.

Munck, Gerardo. 2003. "Measures of Democracy, Governance, and Rule of Law: An Overview of Cross-National Datasets", unpublished manuscript

Murphy, Walter F. 1964. *Elements of Judicial Strategy*. University of Chicago Press.

North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49 (4): 803-32.

North, Douglass C., William Summerhill, and Barry R. Weingast. 2000. "Order, Disorder, and Economic Change: Latin America vs. North America," in Bruce Bueno de Mesquita and Hilton Root, eds., *Governing for Prosperity.* New Haven: Yale University Press.

Russell, Peter H. and David M. O'Brien, eds. 2001. *Judicial Independence in the Age of Democracy: Critical Perspectives from Around the World.* Charlottesville: University of Virgina Press.

Powell, Emilia Justyna, and Jeffrey K. Staton. 2009. "Domestic Judicial Institutions and Human Rights Treaty Violation." *International Studies Quarterly* 53(1): 149-174.

Pozas-Loyo, Andrea, and Julio Rios-Figueroa. 2006. "When and Why 'Law' and 'Reality' Coincide? De Jure and De Facto Judicial Independence in Chile, and Mexico". In Alejandra Rios-Cazares and David Shirk (eds.), *Evaluating Accontability and Transparency in Mexico*, San Diego: U. of San Diego/ Trans-Border Institute.

Przeworski, Adam. 2004. "The Last Instance: Are Institutions the Primary Cause of Economic Development?", *European Journal of Sociology*, vol. XLV, no. 2, pp. 165-188.

Raz, Joseph. 1977. "The Rule of Law and Its Virtue", *Law Quarterly Review*, vol. 93.

Rios-Figueroa, Julio. 2006. "Judicial Independence: Definition, Measurement, and its Effects on Corruption", PhD Dissertation, New York University.

Skanning, Sven-Erik. 2008. "Measuring the Rule of Law". Paper presented at the ECPR Joint Sessions of Workshops, Rennes, April 12-16.

Shklar, Judith. 1998. *Political Thought and Political Thinkers* (ed. Stanley Hoffman). Chicago: The University of Chicago Press.

Simmons, Beth A. 2007. \Untitled Book Project." Working Project. URL:http://www.people.fas.harvard.edu/_bsimmons/ResearchPapers/BookConf/index.html

Souva, Mark, Dale Smith and Shawn Rowan. 2008. "Promoting Trade: The Importance of Market Protecting Institutions." *Journal of Politics* 70 (2): 383-392.

Spriggs, James F. 1996. "The Supreme Court and Federal Administrative Agencies: A Resource-Based Theory and Analysis of Judicial Impact." *American Journal of Political Science* 40 (November): 1122-1151.

Stasavage, David. 2002. "Private Investment and Political Institutions." *Economics and Politics* 14 (March): 41-63.

Tate, C. Neal, Tate, C. Neal, Stacia L. Haynie, Reginald Sheehan, and Donald R. Songer. 2002. "Collaborative Research: Extending a Multi-Country Database and Program of Research." National Science Foundation, Law and Social Science Program, Basic Research Grant.

Tate, C. Neal and Linda Camp Keith. 2007. "Conceptualizing and Operationalizing Judicial Independence Globally." Paper delivered at the Annual Meeting of the American Political Science Association, August 30 – September 1, 2007, Chicago, Illinois.

United Nations. 1985. *Basic Principles on the Independence of the Judiciary.* Office of the High Commissioner of Human Rights. URL: http://www.unhchr.ch/html/menu3/b/h_comp50.htm

Vanberg, Georg. 2005. *The Politics of Constitutional Review in Germany*. New York: Cambridge University Press.

Waldron, Jeremy. 2002. "Is the Rule of Law an Essentially Constested Concept (in Florida)? *Law & Philosophy*, vol. 21, no. 137

World Bank. 2007. *A Decade of Measuring the Quality of Governance. Governance Matters 2007. Worldwide Governance Indicators 1996-2006*. Washington, D.C.

Appendix: Brief descriptions of measures included in the analysis

Political Risk Services (PRS) "Law & Order" measure

The Law and Order measure is part of the Political Risk Rating that includes 12 components and 15 subcomponents covering both political and social attributes. The PRS staff collects political information and financial and economic data, converting these into risk points for each individual risk component. The political risk assessments are made on the basis of subjective analysis of the available information. In the "Law and Order" component, Law and Order are assessed separately; each component can take values from 0 to 3 to a total of 6 joint points together. The Law sub-component is an assessment of the strength and impartiality of the legal system, while the Order sub-component is an assessment of popular observance of the law. Thus, a country can enjoy a high rating – 3 – in terms of its judicial system, but a low rating – 1 – if it suffers from a very high crime rate of if the law is routinely ignored without effective sanction (for example, widespread illegal strikes).

See: http://www.prsgroup.com/ICRG_Methodology.aspx

Wittold Henisz's (2000)measure of "judicial independence"

The existence of an independent judiciary (J=1 or J=0) is determined through the joint existence of a POLITY score on executive constraints of at least 3 (see definition below) and, where data is available, an ICRG score on Law & Order of at least 4 (see definition above). Henisz's measure on judicial independence is available for the period after 1960.

See: http://www-management.wharton.upenn.edu/henisz/

Polity measure on "Constraints on the Executive"

As noted in the POLITY II Codebook (Gurr 1990):

"Operationally, this variable refers to the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities. Such limitations may be imposed by any "accountability groups." In Western democracies these are usually legislatures. Other kinds of accountability groups are the ruling party in a one-party state; councils of nobles or powerful; advisors in monarchies; the military in coup-prone polities; and in many states a strong, independent judiciary. The concern is therefore with the checks and balances between the various parts of the decision-making process. A seven-category scale is used (for examples on evidence used to code countries, see Gurr 1990).

See: http://garnet.acns.fsu.edu/~whmoore/polity/polity.html

Fraser Institute's measure of "Legal structure and security of property rights"

The annual *Economic Freedom of the World* Report uses 42 distinct pieces of data to measure economic freedom in 141 nations. The index measures the degree to which the policies and institutions of countries support economic freedom. According to the Fraser Institute, the cornerstones of economic freedom are personal choice, voluntary exchange, freedom to compete, and security of privately owned property. They measure five broad areas: (1) size of government; (2) legal structure and security of property rights; (3) access to sound money; (4) freedom to trade internationally; and (5) regulation of credit, labor and business. Each component and sub-component is placed on a scale from 0 to 10. The sub-component ratings are averaged to determine each component; the component ratings within each area are averaged

to derive ratings for each of the five areas; in turn, the summary rating is the average of the five area ratings.

The area of "Legal Structure and Security of Property Rights" considers that protection of persons and their rightfully acquired property is a central element of both economic freedom and a civil society. According to the Fraser Institute, the key ingredients of a legal system consistent with economic freedom are: rule of law, security of property rights, an independent judiciary, and an impartial court system. Indicators for this area are assembled from three primary sources: the Political Risk Services (PRS, data generated by staff of the company), the Global Competitiveness Report (GCR, data generated by a survey of firms operating in a given country), and the World Bank's Doing Business data set (DB, data generated by a survey or businessmen).

See: http://www.freetheworld.com/

La Porta et al (2003) measure "Judicial Checks and Balances"

La Porta and his coauthors measure judicial independence and constitutional review. "Judicial independence is of obvious value for securing property and political rights when the government is itself a litigant, as in the takings of property by the state. But judicial independence is also socially valuable in purely private disputes when one of the litigants is politically connected, and the executive wants the court to favor its ally." (2003, 3) "Besides seeking to influence judges, the executive and the legislature would also wish to pursue policies and pass laws that benefit themselves, or democratic majorities, or allied interest groups. Constitutional review is intended to limit these powers. By checking laws against a rigid constitution, a court can limit such self-serving efforts." (2003, 4)

Their measure of "judicial checks and balances" comes from national constitutions, including all 71 countries covered in the Maddex (1995) Encyclopedia of Constitutions, with the exception of transition economies because "their constitutions are rapidly changing". Judicial independence is proxied by looking at the tenure of judges in the highest ordinary court, the tenure of judges in the highest administrative court, and to whether courts have "law making" powers and judicial decisions are constrained by prior judicial decisions. Constitutional review is measured by the degree of rigidity of the constitution, and the extent of judicial review powers.

Feld and Voigt's (2003) measure of *De Jure* and *De Facto* Judicial Independence

Feld & Voigt measure judicial independence *de jure* and judicial independence *de facto*. They focus on the independence of the court with authority to interpret the constitution.

The authors sent a questionnaire to experts (judges, law professors, lawyers, activists of NGOs) asking several quite detailed questions regarding the different components of both of their concepts. The components of judicial independence de jure are (1) whether the highest court is anchored in the constitution; (2) how difficult is to amend the constitution, (3) appointment procedure of judges; (4) their length of tenure, (5) whether there is a fixed retirement age of judges in the court; (6) removal procedures; (7) whether reelection of judges is possible; (8) protection and adequacy of salary of judges; (9) accessibility to the highest court; (10) procedure for allocation of cases in the court; (11) judicial review powers; and (12) transparency of the court.

The components of judicial independence de facto are: (1) average length of tenure; (2) deviation of average length of tenure from de jure prescription; (3) number of judges removed from office; (4) frequency of changes in the number of judges in the court; (5) real salary of judges; (6) real court's

budget; (7) number of constitutional changes in relevant articles; and (8) compliance by other branches on court rulings.

Cingranelli & Richards (2008)

Cingranelli and Richards develop their measure of judicial independence from the United States State Department yearly country reports. Their concept of judicial independence requires that judges be free from control by the government or military. In particular, Cingranelli and Richards ask whether judges can be removed from office, whether there is judicial review, whether judges are free from corruption and whether case outcomes are not influenced by the government.

The measure is ordinal and contains three categories. A state received a score of 2 if there are no limitations on courts, per the rules discussed in the text on page **XX**. A state receives a 1 if there are structural limitations on judicial independence (e.g. the ability of the chief executive or minister of justice to appoint and dismiss judges at will, even if they do not actually do so in the particular year being coded or there is limited corruption or intimidation of the judiciary). A state receives a 0 if there are "active and widespread constraints on the judiciary" (e.g. active government interference in cases or judicial dismissals for political reasons.

See: http://ciri.binghamton.edu/documentation/ciri_coding_guide.pdf

Tate & Keith (2007)
Tate and Keith develop their measure of judicial independence from the United States State Department yearly country reports. The measure is ordinal and falls into three categories. Tate and Keith (2007) provide a comprehensive conceptual review of judicial independence, but do not state an explicit concept of their own. Regarding their concept, Tate and Keith (2007, 4) write:

> While the dimensions of judicial independence conceptualized by each of these scholars and others do not fit together perfectly, we do see a common core across them that allows us to identify two somewhat overlapping sets of distinctions. The first is the distinction between (1) institutional (or collective) independence from the other branches or private and public actors and (2) the independence of the individual judge from the same influences.

Following the Howard and Carey approach, Tate and Keith code states as follows. States receive a 2 if "the judiciary is reported as "generally independent" or is independent in practice with no mention of corruption or outside influence," 1 if "the judiciary is reported to be somewhat independent in practice with reports of (some) pressure from the executive "at times" or with occasional reports of corruption," and 0 if "the judiciary is reported as not being independent in practice; is reported to have significant or high levels of executive influence or interference; or is reported to high levels of corruption. (17)"

Howard & Carey (2004)
Discussed in text on pages 12-13

Apodaca Scale (2004)
Discussed in text on page 14

Clauge *et al* CIM (1999)
Discussed in text on pages 16-17

Bertelsmann Transformation Index (BTI) –Judicial Independence (2008)

The BTI is derived from country expert reports. The aggregate index includes a component for judicial independence. Bertelsmann (2008, 20) write

> An independent judiciary refers fi rst and foremost to how far the courts can interpret and review norms and pursue their own reasoning free from the infl uence of rulers or powerful groups and individuals. This requires a differentiated organization of the legal system, including legal education, jurisprudence, regulated appointment of the judiciary, rational proceedings, professionalism, channels of appeal and court administration.

The BTI judicial independence measure is a scale from 1 (low) to 10 (high). Experts consider whether

> The judiciary is free both from unconstitutional intervention by other institutions and from corruption. There are mechanisms for judicial review of legislative or executive acts. The judiciary is established as a distinct profession and operates relatively independently, but its functions are partially restricted by factors such as corruption and insufficient territorial or functional penetration. The judiciary is institutionally differentiated, but its decisions and doctrine are subordinated to political authorities or severely restricted by functional deficits such as territorial penetration, resources or severe corruption. The judiciary is not institutionally differentiated or is significantly subordinated to religious or political authorities.

See: http://www.bertelsmann-transformation-index.de

Table 1. Three Dimensions of the Rule of Law

Dimension	Related to	Components
Institutional	Contention of governmental arbitrariness	Checks and balances
		Judicial Independence
Individual	Prevention or correction of discriminatory practices in law enforcement	Enforcement of negative rights
		Enforcement of positive rights
Social	Prevention or control of instability and violence	Political violence
		Social violence

Table 2. Summary of Judicial Independence Measures

De Facto Independence Measures	*De Facto Concept*	*Years Available*
Cingranelli & Richards (CR)[#]	Power	1990-2004[#]
Howard & Carey (HC)[#]	Autonomy	1992-1999
Tate & Keith (TK)[#]	Power	1990-2004[#]
Henisz (HEN)	Power	1960-2004
Feld & Voigt (FVF)	Power	2003
Fraser Institute (FRAS)	Unclear	1995; 2001-2005
BTI	Unclear	2006; 2008
Polity Executive Constraints (POL)	Power	1960-2006
PRS Law & Order (PRS)	Power	1984-2004
Clauge et al (CIM)	Power	1960-2000
De Jure Independence Measures		
Feld & Voigt (FVJ)		2003
Apodaca/Keith (APO)		1976-1996
La Porta et al (LAP)		2003

Note. Summarizes the judicial independence measures analyzed here. Descriptions of the measures can be found in the appendix. [#]Measures are made available through the Democracy Assistance Project: http://www.pitt.edu/~politics/democracy/democracy.html.

Table 3. Selected Judicial Independence Scores for selected countries in 2003

	De Facto Measures				*De Jure Measures*	
State	*PRS*	*Fraser*	*Henisz*	*Feld & Voigt*	*La Port et al*	*Feld & Voigt*
Denmark	1	.95	1	.81	1	.78
Germany	.83	.89	1	.8	1	.73
Mexico	.33	.49	0	.71	.33	.80
Pakistan	.50	.23	0	.52	1	.76

Note: All measures are normalized to one that corresponds to the highest degree of "rule of law" or "judicial independence".

Table 4. Correlation Matrix for Judicial Independence Measures

											De jure measures		
	CR	*HC*	*TK*	*HEN*	*POL*	*PRS*	*CIM*	*FVF*	*FRAS*	*BTI*	*FVJ*	*KEI*	*LP*
CR	1.000												
HC	0.616$^{\tau}$	1.000											
TK	0.672$^{\tau}$	0.719$^{\tau}$	1.000										
HEN	0.483$^{\tau}$	0.480$^{\tau}$	0.440$^{\tau}$	1.000									
POL	0.631	0.697	0.602	0.649	1.000								
PRS	0.387	0.449	0.425	0.729	0.374	1.000							
CIM	0.433	0.499	0.477	0.479	0.504	0.513	1.000						
FVF$^{\#}$	0.309	0.254	0.243	0.283	0.203	0.304	0.235	1.000					
FRAS$^{\#}$	0.604	0.624	0.709	0.559	0.382	0.695	0.411	0.544	1.000				
BTI$^{\#}$	0.474	0.552	0.388	0.250	0.602	0.100	0.324	0.445	0.506	1.000			
De jure measures													
FVJ$^{\#}$	0.114	0.183	0.232	-0.198	0.244	-0.165	0.237	0.111	-0.007	0.051	1.000		
KEI	0.271	0.179	0.138	0.306	0.493	0.078	0.256	0.077	-0.129	0.473	0.184	1.000	
LAP$^{\#}$	0.329	0.050	0.323	0.282	0.269	0.252	-0.226	.017	0.356	0.367	0.128	0.216	1.000

Note. Displays correlation coefficients for judicial independence measures derived from all available data. Correlation matrix with the *de jure* category is circled in bold. Shaded coefficients represent cases in which there was no overlapping scores, and so coefficients are derived from the data set used to conduct the settler mortality construct validity test. To match the other Apodaca replication (discussed below) we take the 2006 value for the BTI measure and the 2003 values of Feld & Voigt, Fraser, and La Porta *et al*. All other values come from the 1996 scores of the remaining variables. $^{\#}$Available in 2003 only; $^{\tau}$Kendall's *Tau b*

Table 5. Apodaca Replication: Judicial independence results across twelve political terror models

De facto Measures	*Coefficient*	*RSE*	*p value*	*N*
Cingranelli & Richards				
Partial	-0.079	0.057	0.172	126
Full	-0.480	0.852	<0.000	
Howard & Carey				
Partial	-0.040	0.059	0.502	128
Full	-0.528	0.068	<0.000	
Tate & Keith				
Partial	-0.108	0.057	0.059	128
Full	-0.451	0.074	<0.001	
Henisz	-0.291	0.071	<0.001	120
Feld & Voigt	-0.223	0.284	0.437	50
Fraser	-0.104	0.015	<0.001	85
Polity	-0.108	0.045	0.019	121
PRS	-0.208	0.295	<0.001	94
CIM	-0.581	0.182	0.002	123
De jure Measures				
La Porta et al	-0.245	0.217	0.265	57
Apodaca	0.012	0.009	0.186	123
Feld & Voigt	0.368	0.411	0.375	58

Note. This table displays the OLS estimates with White/Huber standard errors for twelve identical models of human rights violation as described in Apodaca (2004). The only change across the models involves substituting in a new judicial independence measure. All but one coefficient is negative across the models and all but the *de facto* Feld & Voigt measure reaches statistical significance. Full results are available upon request.

Table 6. Relationship between 19th Century Settler Mortality Rate and late 20th Century Judicial Independence

	Settler Mortality Effects			
De facto Measures	*Coefficient*	*RSE*	*p value*	*N*
Cingranelli & Richards[*]	-0.930	0.227	<0.001	58
Howard & Carey[*]	-0.888	0.276	0.001	60
Tate & Keith[*]	-0.804	0.231	<0.001	60
Henisz[**]	-1.93	0.374	<0.001	58
Feld & Voigt	-0.092	0.037	0.022	21
Fraser	-0.878	0.265	0.002	47
Polity	-0.774	0.193	<0.001	56
PRS	-0.510	0.115	<0.001	59
CIM	-0.067	0.010	<0.001	60
De jure Measures				
La Porta et al	-0.026	0.045	0.569	28
Apodaca	-0.332	0.326	0.314	57
Feld & Voigt	-0.077	0.016	<0.001	25

Note. This table displays coefficients from regressing twelve judicial independence measures on the log of the Acemoglu, Johnson and Robinson (2001) measure of 19th century settler mortality. The first three coefficients were estimated via ordered logit, the fourth via logit, and the remainder via ordinary least squares. *[*]Ordered logit model; [**]Logit model.*

Table 7. Simple Non-random Patterns of Missingness in Judicial Independence Measures

	GDP/cap *lnGDP/cap*		
De facto Measures	*Not Missing*	*Missing*	*Difference*
Cingranelli & Richards	5926	4496	1429***
	7.49	7.48	0.005
Howard & Carey	5410	4534	876
	7.46	7.32	0.13
*Tate & Keith**	5382	14985	-9603***
	7.47	8.71	-1.24***
Henisz	5451	3020	2421***
	7.46	7.34	0.12*
Feld & Voigt	9927	4428	5499***
	8.41	7.21	1.21***
Fraser	9856	2423	7434***
	8.25	6.91	1.34***
Polity	5464	4169	1295***
	7.46	7.44	0.016
PRS	6988	2188	4800***
	7.76	6.92	0.84***
CIM	5228	3134	2094
	7.50	7.01	0.47***
BTI	2479	13607	-11129***
	7.08	8.65	-1.58***
De jure Measures			
La Porta et al	10531	3563	6967***
	8.19	7.23	0.95***
Apodaca	5287	4491	886**
	7.46	7.62	-0.16*
Feld & Voigt	9142	4434	4709**
	8.29	7.18	1.11***

Note. This table displays averages for GDP/cap and its natural log between states over which the judicial independence measures are missing and not. We report the results of two tailed t-tests for each difference.

Table 8. Relationships Between Measures by Economic Development

	Take & Keith Correlation Coefficient	
De facto measures	Below Median GDP/cap	Above Median GDP/cap
Cingranelli & Richards$^{\tau}$	0.46	0.73
Howard & Carey$^{\tau}$	0.65	0.70
Feld & Voigt	-0.10	0.26
Polity	0.30	0.68
PRS	0.03	0.40
CIM	0.30	0.40
Henisz$^{\tau}$	0.19	0.46
Fraser	0.52	0.65

Note. This table displays correlation coefficients for *de facto* independence measures below and above median GDP (US1500). $^{\tau}$Row displays Kendall's *Tau b.*