

## **An Evaluation of Cross-National Measures of Judicial Independence**

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

December 28, 2009

### **Abstract**

Judicial independence plays an increasingly important part in theoretical and empirical analysis on substantive problems in the social sciences. Consequently, a number of measures aiming to capture subtle varieties of this concept have proliferated over the past decade. We provide a conceptual map of judicial independence and evaluate the content, construct, and convergent validity of thirteen cross-national measures. We find support for the validity of a number of *de facto* and less support for *de jure* judicial independence measures. We highlight a significant missing data problem in these data and suggest how this problem is compounded by common practices used to demonstrate robustness.

## **Introduction**

Over the last quarter century, scholars have converged on an approach to institutional reform, in which judicial independence is viewed as critical to the promotion and maintenance of many aspects of human welfare. Modern institutional research suggests that independent judiciaries, which constrain arbitrary state power, ensure that state promises to respect individual rights are perceived credible (e.g. North and Weingast 1989). In turn, credibility breeds efficient investment, state solvency, growth and development (e.g. Acemoglu, Johnson and Robinson 2001; Barro 1997; Frye 2004). There is also evidence that independent courts help stabilize democratic regimes, by providing a needed monitoring function for elite coalitions, by coordinating public beliefs about what constitutes a violation of fundamental rules and by generally lowering the stakes of holding power (North, Summerhill and Weingast 2001; Reenock, Staton, and Radean 2009; Weingast 1997). Researchers also have suggested that judicial independence is critical to the expansion and protection of human rights. Indeed, an independent judiciary is commonly considered a necessary condition for a rights revolution (Epp 1998). More generally, there is empirical support for an inverse relationship between judicial independence and human rights violations, including significant problems like state-sponsored torture (Cross 1998; Hathaway 2007; Keith 2002; Powell and Staton 2008).

Beyond these appealing empirics, judicial independence also is regarded well on normative grounds. Most obviously, it is considered a fundamental component of the rule of law (e.g. Raz 1977, 198). Although there is some disagreement over what precisely a thick definition of the rule of law should entail (Waldron 2002), scholars commonly identify three general dimensions: an institutional dimension linked to the prevention of governmental arbitrariness, an individual dimension linked to the prevention or correction of discriminatory practices in law

enforcement, and a social dimension related to social stability (Shklar 1998, 37). For most, the institutional dimension is absolutely critical. 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.” In classic institutional theory, dividing the state into concurrent and mutually dependent powers has been thought crucial to precluding arbitrariness and promoting liberty (Montesquieu 1997; Madison et al 1961); and, an independent judiciary is a fundamental mechanism for advancing this goal (e.g. Montesquieu 1997, Book II). Thus, judicial independence serves as a vital building block for the first rule of law dimension.

For these good positive and normative reasons, a network of scholars, NGOs, states, and international organizations has promoted vigorously the construction of judicial independence around the globe, and more broadly, the rule of law (see Carothers 2006). In order to validate recommendations, to track reform performance, and no doubt because of the concept’s importance to empirical scholarship, the number of judicial independence measures has increased tremendously in recent years. We have identified thirteen cross-national measures of the concept, which cover a wide array of states and a substantial number of years. Clearly, we do not want for data, and as we accumulate more and more information, we should find ourselves in a better position to test theoretical claims about judicial independence and to track legal reform efforts. In practice, however, we lack systematic information on the quality of these measures. For this reason, it is unclear which, if any, are appropriate for particular research projects.

Alarming, scholars have expressed serious concerns about the validity and reliability of cross-national judicial independence scores. Staasavage (2002, 51 fn. 17) has gone so far as to exclude the judiciary in his analysis on the effects of political institutions on private investment because “no accurate cross-country data are available to determine when and where the judiciary

acts as a veto player with respect to policies which matter for investors.” Scholars also have raised concerns about the varied, potentially unreliable, and likely biased processes by which the measures are generated, in particular in the case of subjective measures (be they derived from surveys of firms, entrepreneurs, experts, or citizens). One noted analyst has put the concern as follows: “Because international experts often tended to be businessmen, theirs was a one-sided view and measure” (Hammergren 2006, 14).

These concerns are troubling in ways directly related to basic research on judicial independence and its related policy reform implications, but they are also related to broader concerns with the rule of law. Quite obviously, if our measures are invalid indicators of our concept, it is unclear what to make of our voluminous claims regarding the substantial benefits of building good judiciaries. And surely, we would not want to advocate fundamental legal reforms based on shoddy science, especially in a world of limited resources. But what is more, if our judicial independence measures are invalid, by implication, it is deeply unclear what to make of our rule of law measures, which require scholars to aggregate information about judicial independence along with other pieces of the multidimensional construct. If a measure of one of many underlying dimensions is invalid, then the best we can hope for is that the problem only introduces noise into the scale, rendering an aggregated rule of law measure less efficient. This is the best-case scenario, and others, in which we introduce bias, are certainly plausible. For these reasons, whether we are interested in judicial independence itself, or in the broader concept of the rule of law, we have reason to evaluate the validity of our judicial independence measures. Such an evaluation is our goal.

We begin our analysis by reviewing distinct concepts of judicial independence. We then report on the availability of judicial independence measures designed to capture these concepts

over time and across the globe. Third, we present a number of validation tests for each of the measures in the study. We conclude by drawing implications for future research.

The analysis revealed information of three types. First, we observed features of the measurement landscape, which confirmed what we believed we knew well prior to conducting the study. There is a consensus over a small set of conceptual definitions of judicial independence, and empirical scholars are attempting to measure those concepts. For this reason, though with some important caveats, it should not prove difficult to identify measures of favored concepts for years and states useful to a number of research designs. Our study also revealed information, about which we were confident that the field did not know well. We will present evidence suggesting that measures of *de facto* judicial independence are validly measuring concepts of interest. At the very least, there is reason to question a strategy of excluding a judicial independence measure from a study on grounds of invalidity. Yet, there is far from complete agreement among the many indicators, and a natural implication of this result is that scholars ought to be especially careful to evaluate the robustness of their findings to alternative measurement choices. Since measures are easily obtained, this is not an onerous task. Finally, and perhaps most importantly, our study also revealed information, which frankly, we did not know that we did not know. As we discuss below, we find evidence of related patterns of convergence and missing data, which raise questions of bias and complicate efforts to conduct the robustness analysis that is so clearly needed. In short, our measures agree more among developed countries than among developing countries, and missingness is frequently not independent of development. Not only are our estimates plausibly biased in the presence of unaddressed non-random missingness, but a naïve robustness analysis strategy, one that ignores missingness patterns, is likely to result in over-confident beliefs in the robustness of results to

alternative measures. Since there are a number of techniques available to deal with missing data, this is a correctable problem. But as is true almost always when practice meets principle, solving this problem in common research designs will likely not prove trivial.

## **Concepts of Judicial Independence**

Important conceptual differences characterize the literature on judicial independence (Burbank and Friedman 2002; Russell and O'Brien 2001)<sup>1</sup>. Still, we can distinguish 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 to resolve cases in particular ways.<sup>2</sup> A judge is independent when the output of judicial processes reflects underlying 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 78<sup>th</sup> 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 these decisions are enforced in practice (Cameron 2002). Judicial

---

<sup>1</sup> 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 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.

<sup>2</sup> 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.

independence in this sense means *judicial power*. What judges think is what they produce, and what they produce controls the outcomes of legal conflicts.

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 involves the difference between *de facto* and *de jure* concepts (e.g. Feld and Voigt 2003). The latter deals with formal rules designed to insulate judges from undue pressures, either from outside the judiciary or from within. Institutions like fixed tenure, multilateral appointment procedures, budgetary autonomy, and judicial councils are thought to provide such insulation and thus influence judicial behavior by promoting autonomy or power (e.g. Rosenberg 1991). It is certainly possible that rules of judicial insulation enhance either of the judicial independence concepts, but these are theoretical relationships between distinct concepts that need to be theoretically motivated and tested. It is probably preferable to call these rules what they are: institutions of judicial insulation. That said, since the literature persists in calling them *de jure* judicial independence, we will do so the same here.

With respect to testing claims about the effects of judicial independence on outcomes of interest, there are extremely good reasons to have access to measures that indicate each of these concepts. The need for *de jure* and *de facto* measures is transparent. But it can be useful to have access to measures of the two *de facto* concepts, as well. For example, an argument in which independent courts serve as veto players assumes that decisions are final (Ferejohn and Shipan 1990), and arguments in which independent courts enforce international agreements must assume compliance (Hathaway 2007). These types of arguments require a measure of power. Other arguments do not. Consider a model like Carrubba's (2005), in which independent courts operate as monitoring devices, which assist in the management of cooperative agreements among

political elites (e.g. major political coalitions, states in the international system, sub-state units in a federation). Independent courts, on this account, facilitate the maintenance of cooperative agreements by helping the parties more efficiently enforce their own deals. Here, judicial independence, in the sense of autonomy from the parties, is sufficient to drive the theoretical relationship Carrubba uncovers. Other actors in the model endogenously provide enforcement of judicial decisions, so that the implementation problem is not of interest. However, if judges fail to provide independent information on the extent to which parties carry out their obligations, the mechanism falls apart. A precise test of this argument would only require a measure of autonomy.

### **Data Availability**

We review thirteen cross-national measures of judicial independence, which social scientists have used in empirical research. Figure 1 provides information on the availability of the measures over time, from 1960 to 2008. It divides the measures into categories of *de facto* and *de jure* independence. Within the category of *de facto* independence, we have indicated (whenever it was clear) whether the researchers responsible for the measure intended to capture an autonomy or power concept. The appendix contains detailed information on the construction of the measures summarized in the figure.

[Figure 1]

As is clear from the table, far more research teams have attempted to measure *de facto* independence than *de jure* independence; and, within the former category, scholars largely seem to have targeted the power concept. Importantly, there are only two measures available that provide data prior to the 1980s, and neither of them is unambiguously designed to capture judicial independence only (see appendix). The Polity IV measure captures, in part, legislative

independence. The PRS measure aims to measure order in society, as well as the quality of a state's legal system. The Henisz measure is derived entirely from Polity IV and PRS, so that it inherits the composite nature of both. And finally, the Clague *et al* measure is a proxy for judicial independence, which is more broadly targeted at the suite of institutions that protect property rights. Not shown in the table but also related to availability is the number of countries covered by each measure. The measures described in the figure vary from a sample of 71 countries (Feld and Voigt's measure) to a sample of 163 countries (Polity IV measure).

The bottom line is that scholars have available to them an array of measures of judicial independence, which allow for time series and cross-sectional analysis. Two caveats are worth noting. First, the measures that are available over a decent time series are proxies or composites; however, as we will show below, these measures are positively correlated with those that directly and uniquely target the judiciary. Second, nearly all of the *de facto* measures seem to be aiming at the power concept. Yet, since the power concept itself subsumes autonomy, we would suggest that scholars who are looking for an autonomy measure could plausibly make use of the power measures.

## **Validation Analysis**

We now turn to an analysis of the validity of the thirteen measures we have identified. In order, we provide analyses of content, convergence and construct validity.

### **Content Validation**

A systematic account of content validity would require an evaluation of the extent to which each measure includes information, which ultimately reflects well the theoretical concept of interest. Generally speaking, our view is that the measures do, in fact, contain appropriate content, but there are a number of examples of curious content choices. Ultimately, this is a subjective

exercise, as there is always something to quarrel with when considering the content of particular measures. Rather than review the content of each measure, we identify salient and representative examples of indicators of each sub-concept of independence. We also suggest a few curious choices of content and discuss a problem associated with a plausible data generating process for judicial independence, which nearly all of the research teams seem to disregard.

The Howard and Carey (2004) measure offers a good 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.” This measure, which is derived from United States State Department country reports, is a three-category ordinal scale, 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), who also make use of the State Department’s country reports to generate their measure, 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 certainly possible for two courts to be equally the “authors of their opinions” while the first provides searching review of due process allegations in the criminal law and the second is instead highly deferential in this context.

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 IV Executive Constraints (Polity IV) measure and the Political Risk Service’s Law and Order (PRS) measure. The Polity IV measure indicates the extent to which a state’s chief executive faces constraints on its decision-making process. According to the team, 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 IV and PRS scores, 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. Specifically, since the PRS measure captures not just law, but order, it returns information on the first and third rule of law dimensions. A related practical concern is that these dimensions might be theoretically connected, yet we could not reasonably use the PRS measure to evaluate if this is so. Clearly, it would not be advisable to test a hypothesis linking judicial independence to social order using the PRS measure, precisely because order itself is built into the indicator.

The chief *de jure* measures include Feld and Voigt's survey-based score and Keith's (2002) constitutional indicators of judicial independence, which have been turned into an additive index by Apodaca (2004). Feld and Voigt produce an interval level measure of independence, derived from surveys of country experts. The score 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 only for a limited set of states. 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 from seven items, though Tate and Keith (2007, 23) have questioned the scale's validity, noting that at least two factors capture the underlying conceptual space over which these measures vary.

### **Hybrid Measures**

Although it seems that the *de facto* measures are reasonably targeted at *de facto* concepts, it is not uncommon for these measures to include institutional components, as well. 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. 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.

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. In addition, even though *de jure* indicators are included there is no consideration of the theoretical relations between them, and these rules may be related. For instance, we might offset concerns over relatively short judicial tenure with non-political appointment institutions.

### **Judicial Strategy**

Whether one adopts the power or autonomy *de facto* concept of independence it is worth considering whether and how our measurement strategies address the behavioral implications of strategic models of judicial decision-making. 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 Simmons alludes to is anticipated by common models of judicial-government interaction (e.g. Helmke 2005; Vanberg 2005), and has received considerable empirical support in studies around the globe (Ginsburg 2003; Helmke 2005; Herron and Randazzo 2003). 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.

Once we recognize that judicial decision-making can be strategic, the core measurement challenge is that direct observation of judicial 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 and the extent to which judicial decisions are properly implemented.

For a worldwide sample, comparative judicial politics has yet to produce such data.<sup>3</sup> 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.

One possible solution involves seeking a measure of the behavioral consequences of judicial independence, which are not the subject of analysis. Clague et al's (1999) *Contract Intensive Money* (CIM), which returns "the ratio of non-currency money to the total money supply" (Clague et al. 1999, 188), provides one possibility. 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 extent 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

---

<sup>3</sup> 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.

measures of the concept. 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 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 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 Validation**

Convergent validation considers whether different measures of the same concept return similar information about that concept (Adcock and Collier 2001, 540). Ideally, we would like benchmark indicators of each judicial independence concept, against which we can simultaneously compare our measures. Unfortunately, there is nothing close to a benchmark of either *de facto* or *de jure* judicial independence. In such a setting we can still evaluate the extent to which the measures are associated with each other, which should be true if they are targeting the same concept. Critically, and especially since the power measures should also capture autonomy, we should observe reasonable convergence among all of the *de facto* measures. And we should see that the *de jure* measures are related to one another, as well. Whether we also see

a correlation between *de facto* and *de jure* measures is of less interest, because we are not testing a model of *de facto* judicial independence, and the right theoretical connection between rules and behavior may be highly conditional or non-linear. Figure 2 displays correlation coefficients for each of our measures, first with respect to the Polity IV executive constraints measure and then against the Tate & Keith measure.<sup>4</sup> The figure suggests some cautiously optimistic results for scholars of *de facto* independence. The correlations between the *de facto* measures are all positive, and some are quite large. Importantly, however, and as Haggard *et al* (2008) find, there is considerable variation in the degree of association; some measures seem are only weakly correlated.

[Figure 2]

The relationship between the *de jure* and *de facto* measures is tenuous at best. With one exception, in the left panel of Figure 2, correlations between the *de jure* and *de facto* measures are relatively small. From a measurement validity perspective, this is actually a positive result, as we are looking for evidence that the measures are tapping into distinct concepts. From a theoretical perspective, the result may reinforce the point that formal institutions of judicial insulation are not necessarily related to actual judicial independence (see Herron and Randazzo 2003). But they also raise a 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 more effective in particular political contexts (see Pozas-Loyo and Rios Figueroa 2006).

Figure 3 suggests a different concern. The *de jure* measures are not well related to each

---

<sup>4</sup> Results are expressed in Pearson's  $r$ , though some of our measures are ordinal. We have also calculated the Tau b coefficients for each of the ordinal measures, and the results are unaffected. A complete correlation matrix is available, but the results are quite general.

other. Even the smallest correlation between all of the *de facto* measures and the Polity IV indicator in this figure is larger than the largest correlation between *La Porta et al* and the other *de jure* measures. Basically, the *de jure* research teams have focused on distinct institutional features and there is no evidence that institutional designers pick institutions of judicial insulation as if they came in ready-made bundles. If they did, then focusing on different pieces of a state's complete set of insulating rules would result in roughly the same measure of *de jure* independence. This is not what we observe.

Although there is no benchmark judicial independence measure against which we can compare the indicators under analysis, there is reasonable convergence among the *de facto* measures. Yet, we should not conclude that these measures are perfect substitutes for each other. They seem to return related, but not identical information. We discuss a natural implication of this result below. The same cannot be said for the *de jure* measures. Under the convergence criterion, at least, it appears that these measures are tapping into different concepts of rules that promote judicial insulation.

### **Construct Validation**

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

### **Human Rights**

We begin by replicating Apodaca's (2004) study of state human rights behavior, using twelve of our thirteen measures.<sup>5</sup> 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

---

<sup>5</sup> Due to data availability, it was not possible to estimate the model using the BTI measure.

incentives necessary for courts to constrain governments from violating fundamental human rights protections. Although the precise theoretical arguments linking *de facto* judicial independence to human rights outcomes 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 the average of two scales of political terror (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 score, which increases in the level 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. The research design is cross-section, yet believing that there is a lagged relationship between her independent variables and her outcome variable, Apodaca regresses the 1998 average political terror scale score (PTS) on 1996 values for the independent variables.<sup>6</sup> 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 1 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

---

<sup>6</sup> 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 index 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 that there are lagged effects of growth on human rights practices, the effect of conflict on human rights abuses seems highly likely to be contemporaneous.

two-tailed hypothesis test, and the sample size. Full results are available upon request, but reflect well the findings Apodaca describes in her paper.<sup>7</sup> 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. This suggests that the intermediate category provides little information, which is not already captured by the two ends of the scale.

The Keith *de jure* scale, which Apodaca used and with which she concluded that *de jure* independence deters human rights abuses, did not reach statistical significance in our replication. The coefficient is actually positive. This was initially surprising, but Apodaca notes that she excluded observations for “developed” states. Unfortunately, we do not know how “developed” is defined operationally, so it is impossible to replicate the model exactly. We recovered the negative and significant *de jure* Apodaca result when restricting our sample to states with average GDP below US\$1000, a level well below the median. For the measures not available in 1996, we again find support with the Fraser judicial independence measure, yet neither the Feld and Voigt nor the La Porta *et al* measures produce significant results. What is more, the Feld and Voigt *de jure* coefficient is positive, suggesting that if anything, greater independence on that measure increases human rights violations!

[Table 1]

On balance, the *de facto* independence results are encouraging, and again the *de jure* results are not. Like the convergent validity tests above, it is important to recognize that from a validity perspective, it is encouraging that the *de facto* and *de jure* measures perform differently.

---

<sup>7</sup> Because the publication is in *Judicature*, we do not have access to the actual estimates, but rather just a verbal description.

The models may not suggest much support for an unconditional, linear *de jure* independence effect on human rights outcomes; however, they do not definitively suggest that *de jure* measures are invalidly measuring the concept. Rather, this finding suggests that we might question what institutional feature might 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.

### **Settler Mortality**

Figure 4 summarizes results from another construct validity test, in which we leverage an instrument for “good institutions,” developed by Acemoglu, Johnson and Robinson (AJR 2001). AJR set out to address the obvious endogeneity concern in rule of law models of economic development. AJR instrument for institutional features that allegedly promote development (judicial independence among others) with a measure of the mortality rates among colonial settlers in the 19<sup>th</sup> century. The logic behind the instrument is that where Europeans faced environments that were relatively disease-free and otherwise hospitable, they built institutions that reflected those at home, anticipating significant migration. These institutions were designed to protect rights. Where they faced relatively inhospitable environments, European states did not construct such institutions, as they did not anticipate significant migration. They ran these colonies repressively and did not establish institutions designed to protect rights. These initial choices locked in differences in institutional quality over time.

[Figure 4]

Figure 4 shows six scatter plots of the continuous, *de facto* measures of judicial

independence in the year 2003 (the year for which the cross-sectional measures were available) on the natural log of the settler mortality rate. The figures also show the predicted values from simple, bivariate regressions of judicial independence on the settler mortality rates.<sup>8</sup> Reflecting the convergent validity results above, we observe an almost identical negative relationship between mortality rates and independence across four of the six measures, and all six of the settler mortality estimates are negative and statistically distinguishable from zero.

### **Summary**

So far, our results suggest cautious optimism for using the available measures of judicial independence. Content, convergent, and construct analyses shows that the menu of available measures offers a number of reasonable choices for a variety of research questions. Of course, we should let our theoretical models, within which our concepts operate, guide our measurement choices. Here we echo Collier and Adcock (1999, 539) who argued 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.” More often than not, however, there is not a single best choice of measurement. In this case, given subtle conceptual and measurement differences across the available measures, it is advisable to perform robustness analysis using different measures. As we now discuss, however, this approach is not free from difficulty

### **Missingness and Convergence by Development**

Figure 1 underscores a well-understood feature of judicial independence data – measures are missing for much of the latter half of the 20<sup>th</sup> century. Only the Henisz, Polity, PRS and CIM

---

<sup>8</sup> Each settler mortality coefficient is statistically significant in the models we ran. We do not show results from four logit and ordered logit models of our ordered, *de facto* judicial independence scores. Again, the settler mortality rate estimate was negative and statistically significant.

measures offer anything close to a reasonable time series, and the picture only looks bleaker if we extend the period back to the turn of the century. For this reason, if we wish measures that directly target the judiciary, we are left largely with time invariant measures. In so far as theories of judicial independence commonly predict temporal change, cross-sectional measures are not ideal. 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 by economic development. If this pattern is general, and if scholars do not address it (and we have no evidence that they do), then it may be that *all* estimates of judicial independence effects are subject to the biases associated with various haphazard strategies of imputation or data deletion (King *et al* 2001). The question we address now concerns how bad the missing data problem is in these data.

There are many possible systematic explanations for missingness in these data. We test the simplest mechanism for missingness that we could imagine, suggesting that missingness can be explained by a state's level of economic development, largely because it is more difficult to locate appropriate information about poorer states. We evaluated this possibility in two related ways. First we created a binary missing data indicator for each variable, and then estimated a logistic regression of that indicator on the natural log of GDP/capita. We also conducted difference of means tests for the development measure, across the subsample of state-years for which we have data and the subsample for which we do not.

Table 2 summarizes the results. For each type of test, the measures on the left column reflect those for which we found a statistically significant effect of development on the

probability of a missing value or a difference in the mean level of development across the subsamples. The measures on the top of the table are more likely to be missing for less developed states whereas those on the bottom are more likely to be missing for more developed states. More than half of the measures have missingness patterns that appear dependent on development, and of these measures, they are most likely to be missing for developing countries. Interestingly, none of the measures developed from the State Department reports (Tate & Keith, Cingranelli & Richards, and Howard & Carey) fit this pattern.

[Table 2]

In one sense, this is not particularly alarming. Missingness is a challenge in many research designs, and scholars generally have a responsibility to deal with it as it emerges. But it is important to stress that scholars in our field *are not dealing with it*. Further, Figure 5 suggests that the problem has implications for the robustness analysis suggested by the convergent validity tests above. The figure shows a dot plot of the correlations between the Polity IV executive constraints measure and the other *de facto* scores, above and below the sample's average GDP/capita. For every measure except Feld and Voigt's, correlations are stronger above the average development level in the sample than they are below. Providing this information for all measures is cumbersome, but the results generalize. In some cases, correlations among the *de facto* measures below average GDP are actually negative!<sup>9</sup> Developing countries may not only have lower quality data but also may also experience different paths towards developing the rule of law, which can cause great divergence in expert perceptions (see Weingast 2009).

[Figure 5]

Consider the implications of these results. Recall that the natural implication of having

---

<sup>9</sup> The correlation between Feld and Voigt's *de facto* measure and Tate and Keith's is actually -.41 below average GDP and .34 above. The correlation between the Fraser measure and the CIM is -.37 below average GDP and .54 above.

easy access to distinct measures, which are generally but not perfectly correlated with each other, is to evaluate the extent to which results are robust to alternative indicators. Haggard *et al* (2008) recommend the same strategy in light of the lack of a benchmark measure. While we agree with this suggestion in principle, if we do this without dealing with the missing data problem, we risk biasing our tests toward confirming our inferences.

The reason is as follows. Judicial independence data is missing systematically by level of development for many *de facto* measures. The *de facto* measures are more likely to agree among more developed states. If we fail to address the missing data problem, then we will be more likely to conduct our robustness checks on observations from more developed states, as these are the states for which we are more likely to have information across our measures. These are also the states where our measures are more likely to agree. Consequently, the failure to address this missing data problem in the context of a robustness analysis is that we may be more confident in the robustness of the findings than we should be.

## **Conclusions**

We have centered this paper on the concepts and measures of judicial independence, a crucial variable in several theories of economic development, rights protection, and the rule of law. Although we can quarrel with the content of various judicial independence measures, there is evidence that most are aiming at one of two well-known independence concepts – autonomy and power. We also find 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, but the measures are not always strongly correlated.

Further, 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 the result calls for better theoretical models, which identify how precisely institutions for judicial insulation produce judicial independence.

Finally, we stress that scholars ought to consider the non-random patterns of missing data 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.

## References

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development". *American Economic Review*, vol. 91, 1369-1401.
- 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.
- Barro, Robert J. 1997. *Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press.
- 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. California: Sage, 134-147.
- Carruba, Clifford J. 2005. "Courts and Compliance in International Regulatory Regimes." *Journal of Politics*, vol. 67, no. 3, 699-89
- Carothers, Thomas. 2006. *Promoting the Rule of Law Abroad. In Search of Knowledge*. Washington D.C: Carnegie Endowment for International Peace.
- 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.

- Epp, Charles R. 1998. *The Rights Revolution: Lawyers, Activists and Supreme Courts in Comparative Perspective*. Chicago: University of Chicago Press.
- 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.
- Ferejohn, John, and Charles Shipan. 1990. "Congressional Influence on Bureaucracy". *Journal of Law, Economics, and Organization*, vol. 6, pp. 1-20
- Frye, Timothy. 2004. "Credible Commitment and Property Rights: Evidence from Russia." *American Political Science Review* 98(3):453-466.
- Gibney, Mark, Linda Cornett and Reed Wood. "Political Terror Scale 1976-2006." August 19, 2008, from the Political Terror Scale Web site: <http://www.politicalterror scale.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. <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.
- Hathaway, Oona. 2007. "Why Do Countries Commit to Human Rights Treaties?" *Journal of Conflict Resolution* 8(51):588-621.
- 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.
- 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.

- Howard, Richard M. and Henry F. Carey. 2004. "Is an Independent Judiciary Necessary for Democracy?" *Judicature* (87) 6.
- 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.
- 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.
- 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., 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. 1777. *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.
- 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.

- Reenock, Christopher, Jeffrey K. Staton, and Marius Radean. 2009. *Legal Institutions and the Democratic Order*. Emory University. Typescript.
- 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 Virginia Press.
- 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 Accountability and Transparency in Mexico*, San Diego: U. of San Diego/ Trans-Border Institute.
- Powell, Emilia Justyna, and Jeffrey K. Staton. 2009. "Domestic Judicial Institutions and Human Rights Treaty Violation." *International Studies Quarterly* 53(1): 149-174.
- 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.
- Rosenberg, Gerald. 1991. *The Hollow Hope. Can Courts Bring About Social Change?* Chicago: University of Chicago Press.
- 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.  
<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, 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 & 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. [http://www.unhchr.ch/html/menu3/b/h\\_comp50.htm](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 Contested Concept (in Florida)?" *Law & Philosophy*, vol. 21, no. 137.
- Weingast, Barry. 1997. "The Political Foundations of Democracy and the Rule of Law". *American Political Science Review*, vol. 91, no.2, pp. 245-263.
- Weingast, Barry. 2009. "Why Developing Countries Prove So Resistant to the Rule of Law," in Robert L. Nelson, (ed.), *Global Perspectives on the Rule of Law*. New York: Routledge.

Figure 1: Availability of Judicial Independence Measures

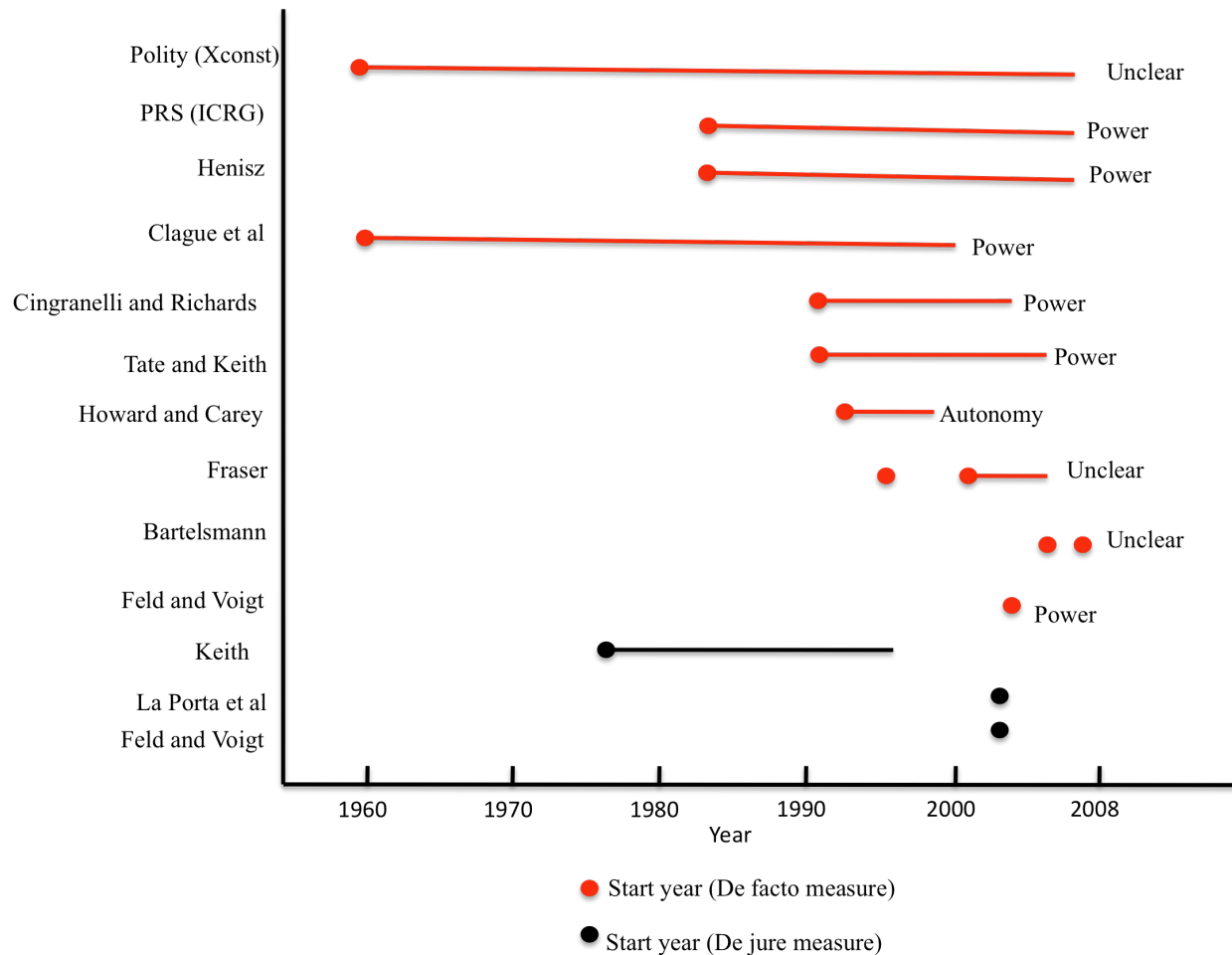


Figure 1. Displays availability of judicial independence measures across time. measures are indicated in red. measure are in black.

Figure 2. Convergence of Judicial Independence Measures

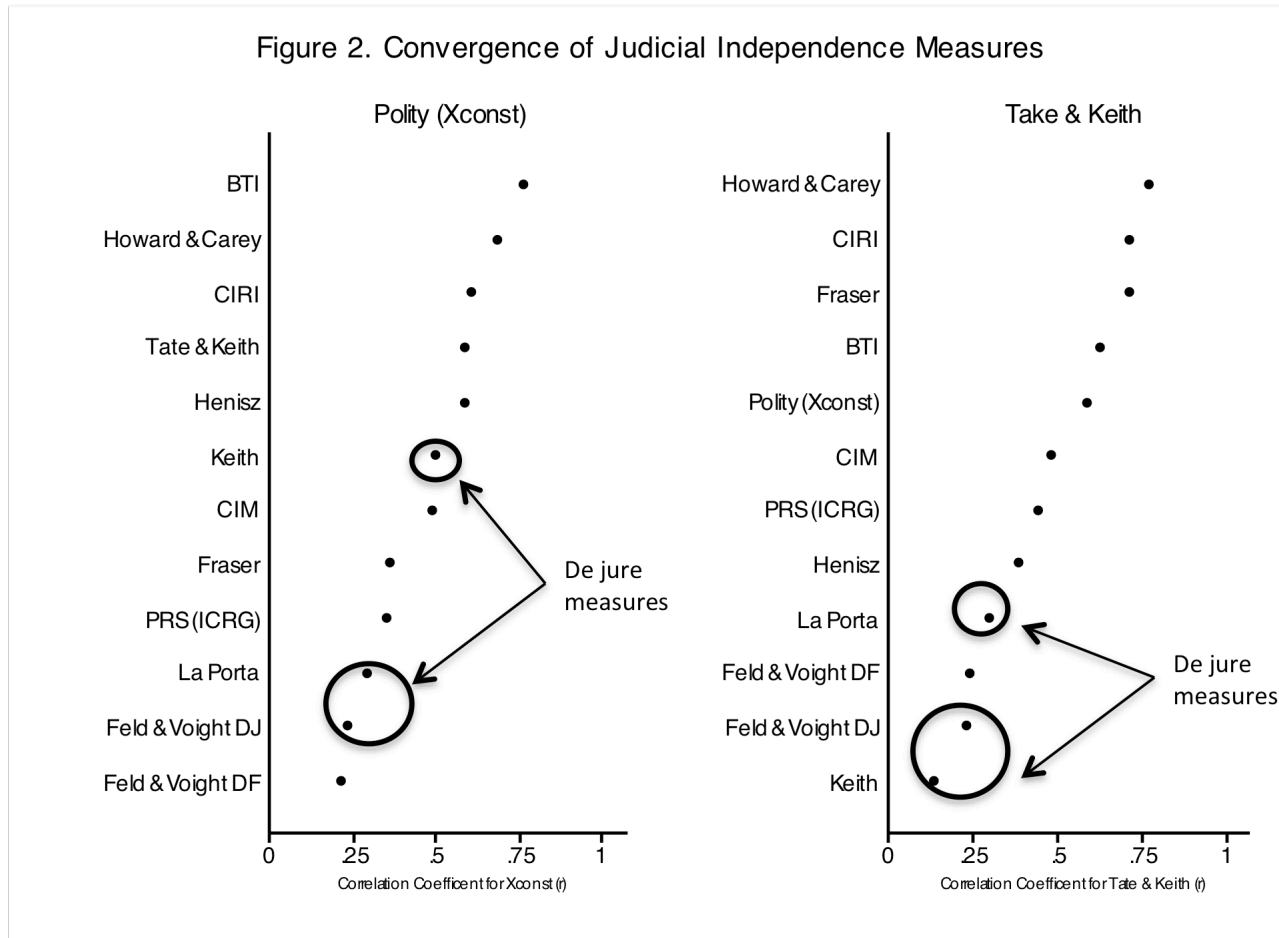


Figure 2. Shows a dot plot of correlation coefficients for each of the measures. Correlations for the Polity IV measure are shown on the left and for Tate & Keith's measure on the right. The results for the three *de jure* measures are circled.

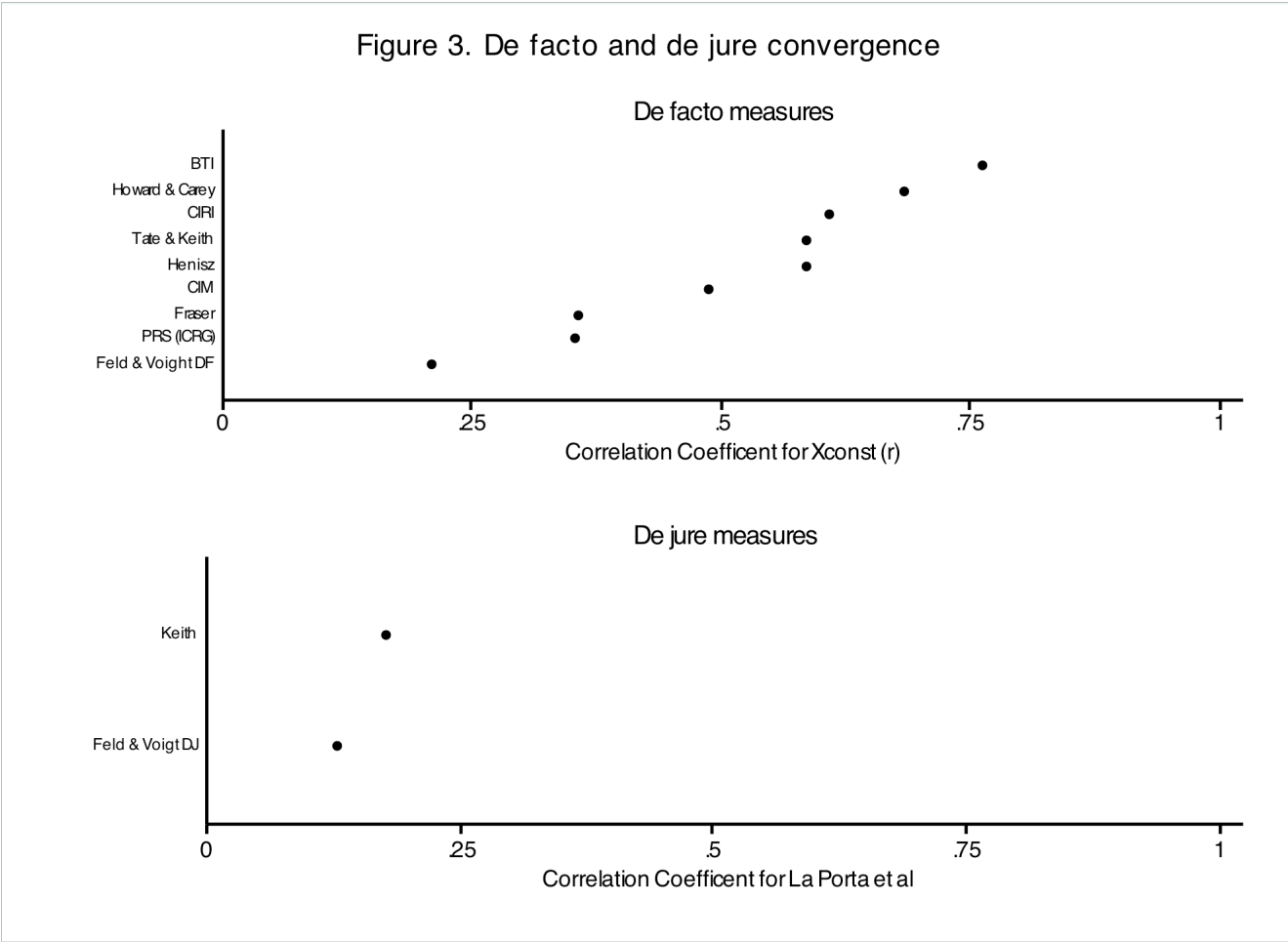


Figure 3. Upper panel displays a dot plot of the correlation coefficients for each *de facto* measure and the Polity IV executive constraints score. The lower panel shows the correlation coefficients for each *de jure* measure with the La Porta *et al* score.

Figure 4. Judicial Independence and Settler Mortality Rate

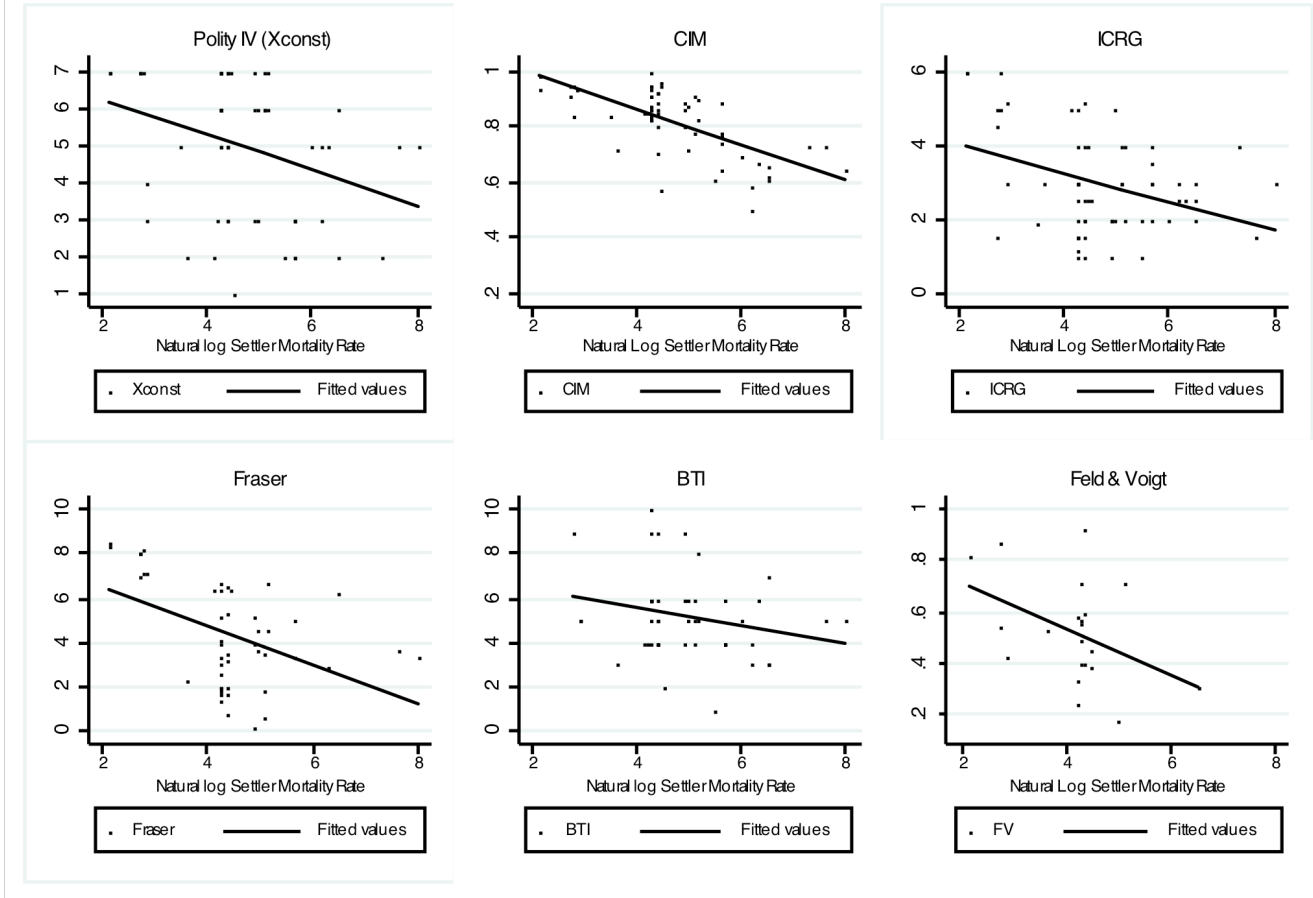


Figure 4. Shows a dot plot of correlation coefficients for each of the measures. Correlations for the Polity IV measure are shown on the left and for Tate & Keith's measure on the right. The results for the three *de jure* measures are circled.

Figure 5. Convergence of Judicial Independence Measures, by Development

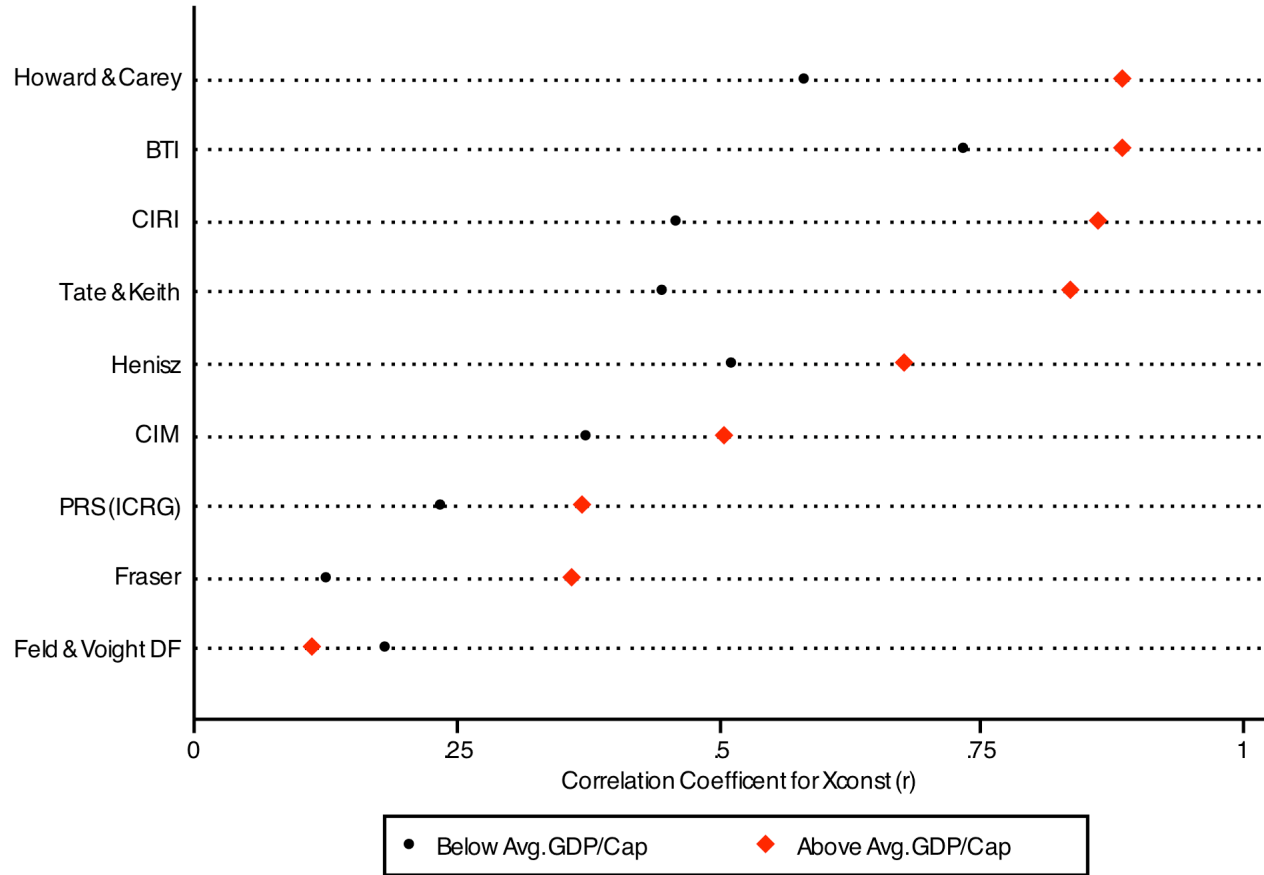


Figure 5. Dot plot of correlation coefficients for each of the *de facto* measures with the Polity IV measure, for states above and below average GDP/capita.

Table 1. Judicial independence results across twelve political terror models

<i>De facto Measures</i>	<i>Coefficient</i>	<i>RSE</i>	<i>p</i>	<i>N</i>
<i>Cingranelli &amp; Richards</i>				
<i>Partial</i>	-0.135	0.049	0.007	138
<i>Full</i>	-0.576	0.068	<0.001	
<i>Howard &amp; Carey</i>				
<i>Partial</i>	-0.082	0.068	0.231	138
<i>Full</i>	-0.605	0.062	<0.001	
<i>Tate &amp; Keith</i>				
<i>Partial</i>	-0.132	0.060	0.030	138
<i>Full</i>	-0.543	0.065	<0.001	
<i>Henisz</i>	-0.302	0.054	<0.001	136
<i>Feld &amp; Voigt</i>	-0.379	0.235	0.114	52
<i>Fraser</i>	-0.104	0.015	<0.001	85
<i>Polity</i>	-0.005	0.002	0.004	131
<i>PRS</i>	-0.236	0.026	<0.001	106
<i>CIM</i>	-0.827	0.175	<0.001	132
<i>De jure Measures</i>				
<i>La Porta et al</i>	-0.290	0.208	0.170	59
<i>Apodaca (Keith)</i>	0.009	0.010	0.328	133
<i>Feld &amp; Voigt</i>	0.219	0.399	0.586	60

*Table 1.* 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 2. Missing Data by Level of Development

	<i>Is Missingness Dependent on GDP/cap?</i>			
	Regression		Difference of Means	
	Yes	No	Yes	No
Missing for low GDP/cap	<i>Feld &amp; Voigt de facto</i> <i>Feld &amp; Voigt de jure</i> <i>Fraser</i> <i>La Porta et al</i> <i>PRS (ICRG)</i>	<i>CIM</i> <i>CIRI</i> <i>Howard &amp; Carey</i> <i>Henisz</i> <i>Keith de jure</i> <i>Tate &amp; Keith*</i>	<i>CIM</i> <i>Feld &amp; Voigt de facto</i> <i>Feld &amp; Voigt de jure</i> <i>Fraser</i> <i>La Porta et al</i> <i>PRS (ICRG)</i>	<i>CIRI</i> <i>Howard &amp; Carey</i> <i>Henisz</i> <i>Keith de jure</i> <i>Tate &amp; Keith*</i>
Missing for high GDP/cap	<i>Polity IV</i> <i>BTI</i>		<i>Polity IV</i> <i>Henisz</i> <i>BTI</i>	

Table 2. For each measure, we estimated a logistic regression of a missing data indicator on the natural log of average GDP, clustering errors within the state for the time series measures. For this analysis, measures for which the development coefficient is statistically significant at the .05 level (two-tailed test) are placed on the left side of the table. Those for which the coefficient is negative are in the top row. We also report the results of difference of means tests for GDP/capita among the subsamples of states for which the measures are missing and present.

\*Excluding the United States of America.

## Appendix

### **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 or if the law is routinely ignored without effective sanction (for example, widespread illegal strikes).

See: [http://www.prsgroup.com/ICRG\\_Methodology.aspx](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, a PRS 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 of businessmen).

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

### **La Porta et al (2004) 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, fully independent judiciary if its judiciary satisfies 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.

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](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.

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)**

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.

### **Apodaca-Keith Scale (2004)**

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.

### **Clague et al CIM (1999)**

The *Contract Intensive Money* (CIM) measure created by Clague et al. 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. The *CIM* was conceptualized as a measure of legal protections for property rights. In order to use it as a general measure of *de facto* independence, one must assume that, on average, states that possess judicial institutions that protect property rights are likely to have judicial institutions that protect rights generally.

### **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 first and foremost to how far the courts can interpret and review norms and pursue their own reasoning free from the influence 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>

## **Appendix References**

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>

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

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.

- Gurr, Ted R. 1990. *Polity II: Political Structures And Regime Change, 1800-1986*. Boulder, CO: Center for Comparative Politics, 1989. Ann Arbor, MI: Inter-university Consortium for Political and Social Research
- Henisz, Wittold. 2000. "The Institutional Environment for Economic Growth", *Economics and Politics*, vol. 12, no. 1, March, pp. 1-31.
- Howard, Richard M. and Henry F. Carey. 2004. "Is an Independent Judiciary Necessary for Democracy?" *Judicature* (87) 6.
- 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.
- 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.
- 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.