

Measuring Judicial Independence in the American States: A Latent Variable Approach

Jeremy R. Johnson*

Pennsylvania State University

April 30, 2015

Abstract

Judicial independence and judicial accountability are commonly understood to exist in tension with one another. Many scholars and professionals, among them the American Bar Association, believe that the judiciary should be entirely independent from any outside influence, be it electoral, executive, or legislative. A contrary view is that judges have become too independent, and need to be reigned in by those who can affect control over them, primarily through judicial elections. The core of this disagreement lies with differing understandings of what constitutes independence and accountability. Absent clear conceptualization and measurement of these concepts, much of the normative debate over judicial independence reduces to disagreement over terms. Despite its manifest importance, there is not currently a comprehensive measure of *de jure* judicial independence in American or comparative politics. I develop and use a latent-variable model to score states' *de jure* judicial independence. This measure will be useful in future studies of judicial independence which focus on *de facto* judicial independence.

*All data and code required for replication is available on GitHub at: <https://github.com/jrjohnson0821/stateslatentreplication>

Introduction

For more than two centuries, judicial institutions have played a key role in the operation of democratic systems of government. An independent court “counteracts the logic of ‘winner-takes-all’ where whoever wins the election wins everything. Thanks to the mechanism of constitutional adjudication, the electoral victory is not an ‘all or nothing’ game” (Ferejohn and Pasquino, 2003, 1685). Normatively, courts provide a secondary outlet for minorities to find resolution of legal problems they cannot otherwise achieve through majority representation. How does this change when the court is selected by the majority as well? When those courts are elected alongside the representatives on partisan ballots, do they have the same incentives for majoritarian rule as governors and state legislators?

The ability of judicial institutions to be effective hinges on their ability to rule against the government, as well as make legal decisions free of influence from other actors such as corporations or interest groups. Ferejohn and Pasquino (2003) characterize this as providing a voice against what John Adams referred to as the “tyranny of the majority” (Adams, 1794). The cost to this independence is that the government and citizens at large have a reduced set of options in holding judges accountable for their decisions. If a Member of Congress or state legislator makes decisions or casts votes in opposition to the public’s interest the public can simply vote the representative out of office. With an independent judiciary this option is generally removed, with the exception of impeachment for criminal acts.

Judicial independence is often discussed as a cause of continuity of the rule of law; it has also been used in many studies of comparative politics as a key indicator of regime stability, economic growth, and most prominently protection of human rights (Linzer and

Staton, 2014, 1). Many scholars agree that an independent judiciary is necessary for the protection of both human rights and political rights (Keith, 2002a,b; Howard and Carey, 2004; Russell and O'Brien, 2001).¹ When examining human rights, judicial independence has often been an important concept; some form of judicial independence is incorporated in the U.S. State Department's Human Rights Scores as well as multiple variables in the POLITY IV dataset (Cingranelli and Richards, 2008; Marshall and Jaggers, 2002; Howard and Carey, 2004). To take but one example, Keith (2002b)'s influential study clearly demonstrates a strong relationship between judicial independence and human rights protections. The World Economic Forum's *Global Competitiveness Report* uses judicial independence as a key indicator in the first pillar of their study of the global business climate (Schwab and Sala-i-Martin, 2014).² The World Economic Forum uses increased judicial independence as an indicator of a better environment for companies to do business. When a foreign company is in a dispute with a local government, an independent judiciary increases the chances that the company will be treated fairly in any legal proceeding rather than the court simply rubber-stamping the government's decision. Judicial independence can be a promoter of other societal benefits such as foreign investment and subsequent economic growth, and consolidation of democracy (Ríos-Figueroa, 2006, 9).

Hayo and Voigt (2007) discuss some of the important factors relating to *de jure* judicial independence, which they refer to as formal independence. First on that list is the credibility of the government. Governments must make credible commitments to an independent

¹For a more extensive review of judicial independence in the human rights literature see (Keith, 2002b, Footnote 1).

²The *Global Competitiveness Report* is a perception based survey of business experts for 144 countries. Judicial independence is covered by the question "In your country, to what extent is the judiciary independent from influences of members of government, citizens, or firms?"

judiciary in order to assure its citizens that they will respect the constitutional rights that they are guaranteed. Formal recognition of an independent judiciary is an important part of that credible commitment. And credibility is a large component of the perception-based surveys that are conducted in research on corruption and human rights.

At the same time, much of the previous research into judicial independence lacks consistency and coherence in both methods and normative definitions. “Judicial independence is often cited, but rarely understood,” (Tiede, 2006, 1). Many scholars have used different indicators of judicial independence based on their specific research question. As Ríos-Figueroa and Staton point out, this lack of clarity of measurement makes it difficult to realistically interpret the influence that formal rules have on judicial independence (Ríos-Figueroa and Staton, 2014, 2). For *de facto* judicial independence, this is discussed at length in Ríos-Figueroa and Staton (2014).

As a result of this, I propose a new theory which combined those indicators which have shown to be informative to *de jure* judicial independence in the American states. This thesis also presents a new measure of modeling *de jure* judicial independence in the American States. The thesis proceeds as follows: Section 1 presents a new theory of judicial independence, Section 2 presents the indicators of this theory, Section 3 discusses the methodology used to create the measurement model, Section 3.2 the corresponding measurement model, Section 4 presents the results from the various specifications of the model, Section 5 tests the validity of this measure against other measures, and Section 6 concludes with proposed applications of this measure to other research.

1 A Theory of *De Jure* Judicial Independence

Defining Judicial Independence

Latent concepts are inherently difficult to measure. They can refer to broad and abstract concepts like democracy, respect for human rights, judicial independence, or smaller concepts such as the ability to do math or reading comprehension, commonly used in educational testing (Jackman, 2008; Treier and Jackman, 2008; Schnakenberg and Fariss, 2014; Linzer and Staton, 2014). The first task in attempting to measure a latent concept is to find a common definition.

There have been many definitions put forth for judicial independence. “Judicial independence can and should be defined as among other things the institutional arrangements designed to protect the judiciary from the executive and the public at large,” (Tiede, 2006, 131). Tiede’s is an abstract conceptualization that is difficult to operationalize. McNoll-gast (2006, 108) present a narrower, but equally abstract characterization of judicial independence “as an outcome that emerges from strategic interactions among the judiciary, the legislature, and the executive.” 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.” Linzer and Staton (2014, 4) use a slightly different definition, stating that a judge is independent “in so far as her decisions reflect her evaluation of the legal regard and in so far as those decisions are respected by government officials.”

Ríos-Figueroa (2006) presents judicial independence as a principal-agent relationship;

“a relationship between those who delegate, in contemporary democracies, the politicians that populate the elected branches of government, and the delegates or the judges in this case,” (Ríos-Figueroa, 2006, 6) While Ríos-Figueroa’s study was designed for a cross-national comparison, this definition is well suited for adaptation in the American context. Rather than simply being the politicians in the elected branch of government, this includes delegation from the voters to the judges through direct elections. Ríos-Figueroa (2006, 17) formalizes this association as follows: “a relation between an actor ‘A’ that delegates authority to an actor ‘B’, where the latter is more or less independent of the former depending on how many controls ‘A’ retains over ‘B’.” In the American context this can be viewed as either the voters or the governor’s control over retention of judges. While voters and governors cannot directly control the decisions that a judge makes, they do exercise considerable power over the retention of judges.

As noted above, one of the challenges in studying judicial independence is that researchers do not share a common conceptual definition of independence (Linzer and Staton, 2014). There is some common ground to these definitions however. Most definitions of judicial independence agree that an independent judge should be free to make her own decisions without regard to public opinion or ruling party positions. They also agree that independent judges should not be intimidated by the public or the government as a consequence of any particular ruling.³ How do judicial systems protect judges from these types of threats to their independence? To answer this question we need to further deconstruct the definition of judicial independence.

³As Ríos-Figueroa and Staton (2014, 4) note, parties and governments put pressure on judges during proceedings by filing *amicus curiae* briefs and other methods, however, these are perfectly allowable within the system. These types of pressure are not considered threats or intimidation, in a similar manner that a letter writing campaign would not be considered a threat to a Member of Congress.

Figure 1: Judicial Independence Continuum



The Judicial Independence Continuum

Many of the institutional rules that remove independence serve the opposite function of creating accountability for judges. [Hall \(2007\)](#) defines two premises of judicial accountability. The first is the simple idea that citizens have accountable control over judges through the ballot box, and the second is that challengers are willing to enter the electoral arena and that the electorate will not always give their full support to incumbents ([Hall, 2007](#), 166). As [Burbank](#) puts it, “judicial independence is merely the other side of the coin from judicial accountability,” ([Burbank, 2008](#), 17).

Figure 1 illustrates that rather than being dichotomous or ordinal, independence and accountability can more properly be considered as a continuum. The Judicial Independence Continuum is a single dimension continuum ranging from a totally accountable judiciary on one side to a totally independent judiciary on the other. As a judicial institution moves across the indicators described below in Section 2, it changes position on the Independence Continuum.

The discussion above also makes clear that judicial independence takes two forms: *de jure* and *de facto* ([Feld and Voigt, 2003](#); [Ríos-Figueroa and Staton, 2014](#); [Rosenberg, 1991](#); [Voeten, 2008](#)). *De facto* judicial independence reflects two concepts: first, a judge is independent when her decisions reflect her preferences and second, even if a court is totally independent “that lacking financial or physical means of coercion, courts depend on the assistance of other political authorities to enforce their decisions” ([Ríos-Figueroa](#)

and Staton, 2014, 4). Can a judge with a set term of office reasonably expect to remain in office no matter which way they rule on a particular case? If the answer to that question is yes, the judge enjoys *de facto* independence.

In contrast, Ríos-Figueroa and Staton (2014, 3) define *de jure* judicial independence as “formal rules designed to insulate judges from undue pressure, either from outside the judiciary or within.” These are the institutional protections that allow a judge to be independent from the executive, public, legislature, or superior courts. Examples of these protections include selection and retention methods, term lengths, and docket control.

Previous Studies

Existing efforts to operationalize *de jure* judicial independence have taken on a wide variety of forms. This is unsurprising, as most such measures were developed with specific research goals in mind, and are thus designed to capture particular aspects of independence that were of particular relevance to the study in question. At the same time, all share some common characteristics. First, all such measures are predominantly institutional in nature, in that they rely primarily on characteristics of the judicial institutions themselves as indicators of independence or lack thereof. Second, nearly all are observational (as opposed to interview-or survey-based) and rely on aggregation of varying numbers of partial indicators into a summary scale. Finally, all were developed for the cross-national context. The latter point is important, since indicators appropriate to a cross-national measure may or may not be similarly valid in the context of the U.S. states (and vice-versa).

Choi, Gulati and Posner (2010) developed a theoretical model of judicial independence in the form of a principal-agent relationship. One of the key concepts of independence in

this relationship is the identity of the principal. In this case, the court as a whole but also individual judges act as the agent, and the appointing authority acts as the principal. As [Choi, Gulati and Posner](#) note, independence hinges on who is responsible for the initial appointment but also on who holds the power of reappointment. As the number of principals increases, so does the amount of those who must be satisfied in order for the judge to maintain their position. This principal-agent relationship is the basis for many of the assumptions in this model. [Choi, Gulati and Posner \(2010\)](#)'s choice of indicators for judicial independence is entirely *de facto* however. [Choi, Gulati and Posner \(2010, 296\)](#) put forth a hypothesis based on a collective-action problem. When judges are elected by the public at large, instead of a single principal, there is instead an exponentially larger number, increasing the free-rider problem. One of the important components of a principal-agent relationship is the capacity of the principal to monitor the agent. In elected relationships, there are hundreds or thousands of principals who are given an incentive to shirk their monitoring responsibility by assuming that another citizen will do it, these are called free-riders ([Olson, 2009](#)). By contrast, in appointment states, with a single governor as the principal, there is no free-rider problem. This is only marginally increased when using a commission-based system, and is perhaps even attenuated by the lack of other political considerations of the committee as opposed to the governor. However, in all commission system states, the governor is ultimately responsible for appointing the judges in question. For example, their primary indicator for judicial independence is a ratio of the number of times that a judge writes an opinion that disagrees with their co-partisans.

[Schmidhauser \(1987\)](#) developed an eleven part framework for a legal system based on neo-Weberian concepts. One of the components of this framework is the concept that

judicial institutions are based upon constitutional rather than ordinary statutory authority, judges are constitutionally protected with life tenure, and full establishment and acceptance of judicial review (Schmidhauser, 1987, 46-47). The American states have accepted some of these concepts but clearly not all. Instead, the American states use a mix of selection methods and tenure lengths. Many of the other indicators used in previous studies, such as constitutional interpretation and judicial review, do not have any variation in the American context and are constitutionally protected in most state constitutions.

Table 1 shows a comparison of previous studies of *de jure* judicial independence. Ríos-Figueroa and Staton (2014) examine three measures of *de jure* judicial independence: Feld and Voigt (2003), Keith (2002a), and La Porta et al. (2004). Table 1 shows the commonality of some of these measures such as appointment/retention methods and term lengths. It also shows the lack of congruence in indicators of *de jure* judicial independence shown in previous studies.

As Ríos-Figueroa and Staton (2014) note, there are some issues with the validity in each of these measures. One main concern is that some of these measures are not directly related to independence. Some of these measures are court accessibility, case allocation, right to strike, the four freedoms, pay comparisons, bans on torture, and legal origin (Feld and Voigt, 2003; Keith, 2002a; La Porta et al., 2004; Ríos-Figueroa and Staton, 2014). The second concern is that in Ríos-Figueroa and Staton study many of the *de jure* studies as well as many of the *de facto* studies use indicators that capture both *de jure* and *de facto* independence. La Porta et al. (2004) attempt to measure *de facto* independence by using *de jure* independence as a proxy. However, this is complicated by the fact that even though the institutional arrangements that derive *de jure* independence are observable, *de*

jure independence itself is still a latent concept.

Table 1: Comparison of Indicators of *De Jure* Judicial Independence

Indicators	Feld and Voigt (2003)	Keith (2002a)	La Porta et al. (2004)	Ríos-Figueroa (2006)	Melton and Ginsburg (2014)
Anchored in the Constitution	X				
Amendment Difficulty	X				
Appointment Procedure	X		X	X	X
Tenure	X		X	X	X
Renewable Terms	X				
Salary Insulation	X			X	X
Pay Comparison	X				
Court Accessibility	X				
Case Allocation	X				
Constitutional Interpretation	X		X	X	
Published Decisions	X				
Four Freedoms		X			
Right to Strike		X			
Writ of <i>Habeas Corpus</i>		X			
Public Trial		X			
Fair Trial		X			
Ban on Torture		X			
Case Law			X		
Legal Origin			X		
Number of Courts				X	
Number of Judges				X	
Budget Control				X	
Impeachment				X	X
Promotions				X	
Transfers				X	
Sanctions				X	
Statement of Judicial Independence					X

Melton and Ginsburg (2014) create an index that combines six aspects of judicial independence. This model as well as that of Feld and Voigt (2003) and Keith (2002b) use an additive index. The studies in Table 1 show the wide variation in indicators used, as well as the most basic method of creating a measure. Only two of these measures, appointment procedure and tenure, are used in more than three of the five studies shown.

In the past studies of latent variables were often constructed using additive models. The most widely cited use of this in comparative politics is POLITY IV (Marshall and Jaggers, 2002). This approach has also dominated the study of judicial independence: most if not all previous studies use additive indices (Feld and Voigt, 2003; Keith, 2002a; La Porta et al., 2004) and some still continue to do so today (Melton and Ginsburg, 2014). Following work in other areas, more recent work has adopted an IRT approach (Martin and Quinn, 2002; Treier and Jackman, 2008; Schnakenberg and Fariss, 2014; Linzer and Staton, 2014; Fariss, 2014).

Take the example of selection method. When using an additive scale, the difference between a partisan and non-partisan election are considered to be a one unit increase along the latent trait of judicial independence. That same increase (of one unit) is also the change from a non-partisan election to a system of legislative appointment. In the latter case, there are vastly fewer people involved in the selection process between these two, reducing the amount of campaign funds needed. The same dynamic can be seen between the levels of term length. Moving from a short term (say 1-3 years) to 4-6 years is not the same as moving from a ten year term to a lifetime or quasi-lifetime term.

2 Indicators

As shown in Table 1 the most commonly employed indicators in previous studies of *de jure* judicial independence are appointment procedure and tenure. However, each author codes these slightly differently. [Feld and Voigt \(2003\)](#), for example, use an ordinal scale for judicial tenure, while [Melton and Ginsburg \(2014\)](#) use a dichotomous variable for lifetime terms. The same is true for appointment procedures. In addition, all previous studies were cross-national in nature, and do not allow for the various forms of elections that occur under the American system.⁴

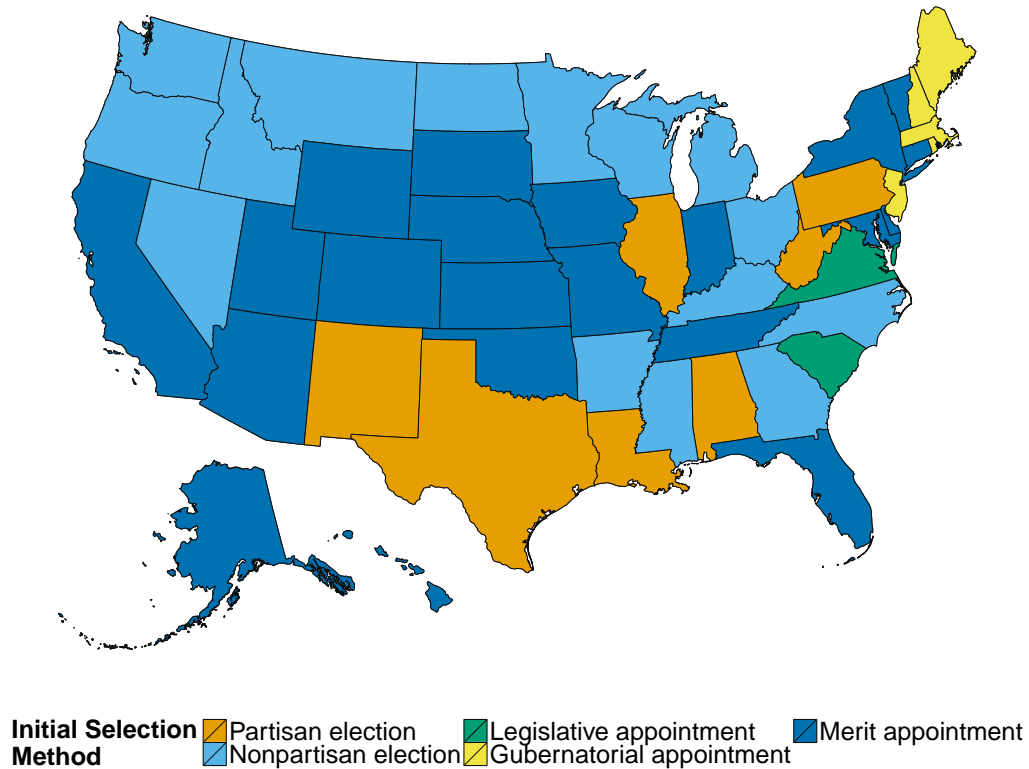
Selection Method

There are five primary selection methods in use in the United States: executive appointment, legislative appointment, merit selection, non-partisan election, and partisan election. Figure 2 shows the current selection methods in use through the United States. Some states use a mix of these methods; for example Michigan and Ohio, both employ systems which blend partisan primaries with nonpartisan general elections.⁵ Under Hall's (2007) theory of judicial accountability, I consider executive and legislative appointments to be the most independent/least accountable of the selection methods. For the majority of states, exec-

⁴[Linzer and Staton \(2014\)](#) create a measurement model that exclusively models *de facto* judicial independence. The measurement model created below is a complimentary addition to that created by [Linzer and Staton](#). This model focuses exclusively on *de jure* judicial independence, but uses a similar statistical model to that of [Linzer and Staton](#).

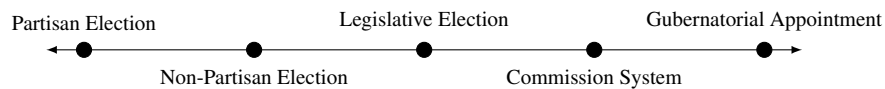
⁵Following normal practice in this field, these states (Ohio and Michigan) are coded as non-partisan ([Canes-Wrone, Clark and Park, 2012](#); [Caldarone, Canes-Wrone and Clark, 2009](#)), however [Tiede \(2006\)](#) does the opposite and codes them as partisan states. [Nelson, Caufield, and Martin \(2013\)](#) discuss this coding decision at length. In accordance with recommendations made in their article, I coded these states both ways and examined for differences, with theoretical justifications for both methods made clear.

Figure 2: Current Judicial Selection Methods



utive appointment was the original method of appointment.⁶ Figure 3 shows how these methods map on to a single dimension continuum with the most accountable method on the left and the most independent method on the right.

Figure 3: Selection Method Independence Continuum



⁶In all executive and legislative appointment states, the governor is the final appointing authority, and judges require a confirmation vote of one or both houses of the state legislature.

The second type of selection system is the commission system or merit selection. The commission system plan came about with the adoption of the Missouri Nonpartisan Court Plan in 1940 (Watson and Downing, 1969). This system was created with the goal of removing partisanship in judicial selection. It created a judicial nominating commission who reviews applications and interviews candidates for judicial posts. The commission consists of three lawyers selected by the Missouri Bar Association, and three citizens selected by the governor as well as the Chief Justice of the Missouri Supreme Court, who serves as the chair of the commission. When a vacancy exists in the courts, the commission submits three names to the Governor who then selects one nominee, who is then confirmed by the State Senate. Once the judge has completed at least one year in office, they stand for retention election. Retention elections are conducted on ballots that are separate from general election ballots and only ask the voter to decide whether or not the judge should be retained. This description of the Missouri Plan is representative of the other twenty states which use similar systems. In most instances, there is no partisan affiliation listed on the ballot (Watson and Downing, 1969). The most substantive difference between commission systems and other systems is the attempt to remove partisanship completely. Appointment, as well as partisan elections, are both acknowledged to be partisan acts, while proponents of commission systems and non-partisan elections claim that these systems eliminate partisanship from judicial selection.

Electoral systems fall broadly into two categories: partisan and non-partisan. Electoral systems are usually considered to be the least independent due to their high accountability (Choi, Gulati and Posner, 2010). These judges are also generally forced to raise money for advertising and other campaign related expenses. Debate exists about the ethical im-

plications of raising this money from businesses, interest groups, and litigators that may appear before these judges.⁷ Elected judges are selected by the electorate in statewide elections.⁸ [Caldarone, Canes-Wrone and Clark \(2009\)](#) put forth the view that the public will elect judges that most closely match their own personal preferences. However, [Debow et al. \(2013\)](#) state that “non-partisan judicial elections do not compare favorably with partisan judicial elections” because the former inhibit the voters ability to make informed decisions.

Retention Method

The creation of retention elections was a hallmark of the merit selection plan that originated in Missouri. Retention elections differ from partisan and non-partisan elections in that there is only one candidate involved in the election. As [Reid](#) summarizes:

“Judicial retention elections are intended to preserve the courts role as an impartial and detached resolver of disputes by ensuring that judges can retain their seats without engaging in the fund raising, politicking, and electioneering that characterize political elections and the political process,” ([Reid, 1999](#), 68).

Retention elections simply ask voters to approve or disapprove of the judge in office. Ballot questions usually appear in the form of: “Shall each of the persons listed be retained in office as Judge of the Appellate Court, First Judicial District?”

⁷For empirical evidence of this see [Gibson \(2008\)](#) and [Hall \(2014\)](#).

⁸While this is true for the majority of states, it is not true for all. Some states such as Louisiana and Illinois, select supreme court justices strictly in districts, while others (such as Tennessee) hold statewide elections but require that the justices be geographically representational ([Bonneau and Hall, 2010](#)).

All but the four states which have life or quasi-life terms employ some manner of retention elections. Some states use retention elections such as those listed above. Others use the same type of partisan or non-partisan election as they use for their initial selection. No state that uses a commission system then requires the judge to run in a contested election.⁹ Also, no state utilizes a partisan election for initial selection and then uses non-partisan elections for retention. Contested elections, whether partisan or non-partisan, are usually considered to be the least independent system for judicial retention (Choi, Gulati and Posner, 2010; American Bar Association, 2003; Canes-Wrone, Clark and Park, 2012). Conversely, judges who do not face any electoral accountability are generally considered the most independent in this context.

Figure 4: Retention Method Independence Continuum

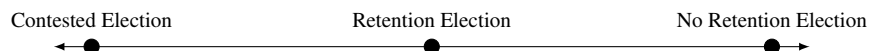


Figure 4 shows how methods of retention map on to the Judicial Independence Continuum. Retention elections are admittedly only a matter of degrees more or less independent than contested elections and there is still some debate on whether retention elections create a more independent judiciary. The ABA defends its position that retention elections are more independent, and recommends them for all states that utilize elections to retain judges (American Bar Association, 2003). Canes-Wrone, Clark and Park (2012) take a divergent position and posit that judges are susceptible to losing their seats based on only one or two salient decisions and also susceptible to attacks from single-issue groups. This

⁹The exception to this is in the case of a judicial vacancy. If a judge resigns, or dies in office, several states have selection mechanisms in place that can differ from the official selection system. For example, in the case of New Mexico “All judicial vacancies are filled by the governor from a list of candidates recommended by a judicial nominating commission. The appointee must then compete in a partisan election at the next general election to serve the remainder of the unexpired term” (*Judicial Selection in the States*, 2014).

occurred in 2010 when three justices on the Iowa Supreme Court were ousted after voting to strike down Iowa's ban on same-sex marriage. Following a prolonged campaign by single-issue groups opposed to same-sex marriage, the judges in question were removed in a retention election ([May, 2013](#)).

Term Lengths

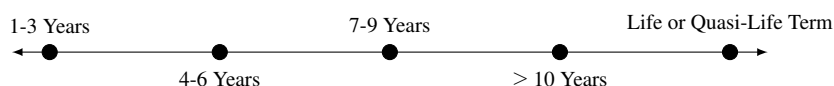
The length of a judicial term is another important factor in determining the level of judicial independence. In many commission system states, seats are vacated by either promotion, retirement, or death. Replacement judges are then appointed to fill out the unexpired terms. The initial term lengths for the new judges in those states range from appointment until the next general election, or from appointment through the remainder of the full term. Subsequent term length is the length of the terms after the initial term. For most elective states this is the same as the initial term length, and for retention states, this is the length of a standard term. Term lengths for initial terms range from one year to fourteen years¹⁰ while subsequent terms range from six years to lifetime terms.¹¹ These term lengths can be compared to U.S. House of Representatives terms which are two years, and to U.S. Senate terms which are six years. In general, the longer that a term is, the more independent choices a judge can make before being held accountable in an election. Figure 5 shows how varying term lengths map on to the Independence Continuum. According to [Ríos-Figueroa \(2006, 31\)](#) “terms must be long enough to reduce the vulnerability of Supreme Court judges. Tenure need not be for life, but if it coincides with that of appoint-

¹⁰While there are four states which have lifetime or quasi-lifetime (age limited terms) only New Jersey requires reappointment after an initial term.

¹¹Many states also have statutory term lengths also utilize a mandatory retirement age.

ing authorities then there is the potential for abuse.”¹² As Hall and others show, most of strategic voting of judges is done in the year prior to an election (Hall, 1987a, 1985; Brace and Boyea, 2008; Canes-Wrone, Clark and Park, 2012).

Figure 5: Term Length Independence Continuum



There are several arguments for the inclusion of term length as an indicator of judicial independence. The ABA views the term length question as a matter of experience in that lawyers who are experienced and qualified will be reluctant to leave successful private practices if they are forced to return to the job market after only a few years (American Bar Association, 2003, 97). The argument against longer terms is that longer terms provide less opportunity for accountability. The ABA recommends an actively enforced judicial discipline system, which will ensure accountability for systems that do not rely on re-selection stating that judges should only be removed for “cause” (American Bar Association, 2003, 103).¹³ The low level of judicial turnover disputes this point of view. In 2012, only 6 of the judges on state supreme courts were not retained. For the years 1964-1998, there were 4,588 retention elections with only 52 judges voted out of office (Aspin, 2000).¹⁴

¹²Ríos-Figueroa (2006) uses a binary coding scheme in his study, coding a term length as independent if it was longer than the appointing authority.

¹³This is in reference to an ethical or criminal violation, rather than a punitive punishment as the result of an opinion. The clearest example of this would be the Iowa case where judges were ousted based on an ideological view, rather than as a breach of ethics or the law.

¹⁴There are several arguments for a connection between longer term lengths and increased judicial independence, as well as for alternatives such as limiting judges to only a single term. A single term would obviate the need for reappointment or campaigning for their next term (Carrington, 1998). Former Presiding Judge of the Texas Court of Criminal Appeals Morris Overstreet has stated “if you don’t have to go back and

Hall's (1987b) study of the Louisiana Supreme Court shows that when faced with electoral accountability, the judges have a tendency to avoid casting votes on unpopular issues. Hall concludes that this system "does not present the opportunity for the unfettered exercise of individual preferences" (Hall, 1987b, 46). Hall also concludes that justices who perceive themselves to have views inconsistent with public opinion, and who desire to keep their jobs, may be hesitant to publish opinions which are against public opinion. Hall states that "whether voter's and opponents are cognizant of the justices behavior or not, certain justices seem to fear the prospect of electoral sanction and consequently alter their behavior" (Hall, 1987a, 1123). While voters knowledge of justices' records is questionable, certainly the judges' actions are not. As Hall shows, judges act as if voters will have knowledge of their records and vote strategically to minimize the electoral losses they may suffer (Hall, 1987a).

Docket Control

The ability of a court to control its own docket is another important factor in examining *de jure* independence. Docket control comes in many forms, ranging from complete discretionary control (as in the case of the U.S. Supreme Court, which can grant or deny any appellate case that it wishes to hear with few exceptions) to state supreme courts such as Wyoming's which have no intermediate appellate courts, and therefore can exercise very little control over its appellate docket. There are also many states which have mandatory appeals on certain subject matter or litigants, and discretionary appeals on other types of

face the voters, you don't have to worry about how they're going to retaliate or if they're going to retaliate" (American Bar Association, 2003). At the same time, this approach has the drawback of returning judges to private practice, which may have an effect on their decisions made while on the bench. No state currently has adopted this practice.

cases. Typically, docket control is determined by statutes passed by the state legislature, similar to the appellate jurisdiction of the U.S. Supreme Court, although some states have their discretion set through the state constitution.¹⁵

Figure 6: Docket Control Independence Continuum



The ability of a court to control its docket has multiple relationships to its effect on independence. The first is that if a court has control over its docket, then the workload of the court can be reduced to only those cases that the court wishes to make a statement on, rather than hearing every case, including those on which there is already clear precedent (Maltzman, Spriggs and Wahlbeck, 2000). Also, reduced workload enables a court to spend more time on those cases that it deems important. While this may seem more related to efficiency than independence, the two are inter-connected in this case. As Squire (2008) notes, the ability to pick and choose cases allows an appellate court to craft their decisions more carefully than a court which has no discretion.

The second influence of docket control on judicial independence stems directly from the first. Not only can judges spend more time on cases under a discretionary docket, but they can pick and choose the cases that they want to issue an opinion on. As Fontana notes, this can be very important to both the influence and the legitimacy of a court (Fontana, 2011). Fontana notes that had the *Brown v. Board of Education* decision been decided ten years earlier, it might not have had the effect that it had, due to a lack of support from

¹⁵Death penalty cases are an exception to this general rule. Due to a Federal requirement that death penalty appeals be heard by the court of last resort, there is no variance across states. In the case of this analysis, this exception will be ignored. For states which utilize complete discretionary docket control except for death penalty cases, they will be coded as having complete discretion.

the President and antipathy from southern states and their representatives in Congress. This view is supported by [Rosenberg \(1991\)](#) in light of the very extensive process to bring the country in to compliance with the Court's ruling in *Brown*. Had this process happened earlier, there might have been no compliance, causing both extended oppression of African-Americans in the South, as well as a severe blow to the legitimacy of the Court. Fontana refers to this as "legitimacy timing" ([Fontana, 2011](#), 627).

Another important aspect of docket control is the effect that it has on a judge's voting preferences. Whether following a strictly legal interpretation, or ideological preferences, a key part of judicial independence is a judge's ability to vote according to his or her preferences. [Songer, Ginn and Sarver](#) note that when courts have discretionary control over their dockets, the judges are much more likely to follow the attitudinal model of voting preferences ([Songer, Ginn and Sarver, 2003](#)). This would lead us to believe that discretionary docket control allows more freedom in the preferences that judges exercise in this case. This is also supported by Hall's study of the Louisiana Supreme Court ([Hall, 1987b,a](#)). Hall finds that in discretionary cases, there is a significant number of dissents, rather than in mandatory cases, which shows much more unanimous decisions ([Hall, 1985](#)). There are several hypotheses for why this is the case, chiefly that under mandatory review, there are many more "easy" cases; that is cases which have a clear precedent and had no complex legal complications. These cases could have been successfully disposed of by the lower courts, but were required to be heard by the state supreme court.

The most important aspect of a court's jurisdictional control comes from their independence from the executive and legislature. This is especially a concern when the government is a party to a case that the court could potentially hear. As [Ríos-Figueroa and](#)

Staton (2014, 6) note, “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 influence from the outcomes of legal conflicts difficult.” They also posit that judicial decision making can appear to be independent if the court removes contentious cases from its docket. The example of same-sex marriage can illustrate this quite clearly. Returning to the gay marriage example, had the Iowa Supreme Court simply chosen to uphold the lower court ruling in *Varnum v. Brien* (763 N.W.2d 862 Iowa 2009), they most likely could have avoided the lengthy campaign that was mounted against the retention of three justices that ultimately resulted in their dismissal.

3 Data

My model uses institutional data on the courts of last resort for the 50 states in the United States. There are 52 courts of last resort in the United States.¹⁶ 48 states have a single court while Texas and Oklahoma have specialized courts for civil and criminal. In this study these court systems will be examined together because they have identical institutional arrangements.

I create two models of *de jure* judicial independence. The first model uses all five indicators described above. The states had a variety of control over their dockets from 1970 through the present. However, between 1940 and 1970 the data is nearly non-existent due to a lack of reliable surveys of state court organizations. The post-1970 data was provided

¹⁶The court of last resort for the District of Columbia is the District of Columbia Court of Appeals. As the District of Columbia is not a state and its courts fall under the federal system, it is omitted from this study. The D.C. Superior Court is the trial court and refers all appeals to the D.C. Court of Appeals. Its judges are appointed by the President and confirmed by the Senate to fifteen year terms.

by the Bureau of Justice Statistics (BJS) and the National Center for State Courts (NCSC) (Rottman, Flango and Lockley, 1993; Rottman et al., 1998; Rottman and Strickland, 2004).

The second model specification is a regime change model, which includes only state-year observations that indicate a change in any of the four indicators used in the primary specification, excluding docket control.

In addition to the data being non-existent, there was quite a bit of variance in the interpretation of docket control based on the individual states. In many of these states, even if there was an “appeal by right,” some states interpreted this as mandatory jurisdiction, while others did not. These discrepancies lead toward a more *de facto* indicator than is being addressed in this paper. With the standardized definitions used in the BJS and NCSC reports, it becomes useful to add this indicator in for a post-1970 model, but as shown below, the wider variance in selection/retention methods and term lengths, still make a study of a longer time span useful, despite the lack of jurisdictional control data.

3.1 Indicators

All of the indicators are ordinal variables, with 0 representing the most independent category and higher values moving towards least independent. Appointment method is range from zero to four, which is the appointment method that is used for a judge’s initial appointment. Initial Term Length ranges from zero to four as well and represents the length of the judge’s initial term. Subsequent Term Length is coded in a similar manner as Initial Term Length and represents the length of the judge’s subsequent terms after their first retention or reappointment.

There is one major difference between these two measures in that zero represents those

Table 2: List of Indicators

Variable	Code
APPT	Appointment Method
0	- Gubernatorial Appointment
1	- Commission System
2	- Legislative Appointment
3	- Non-Partisan Election
4	- Partisan Election
TERM1	Initial Term Length
0	- Lifetime or Quasi-Lifetime Term
1	- ≥ 10 Years
2	- 7-9 Years
3	- 4-6 Years
4	- 1-3 Years
TERM2	Subsequent Term Length
0	- Lifetime Term
1	- ≥ 10 Years
2	- 7-9 Years
3	- 4-6 Years
4	- 1-3 Years
RETELE	Method of Retention
0	- No Retention Election
1	- Retention Election
2	- Contested Retention Election
DOCKET	Docket Control Discretion
0	- Complete Discretion
1	- Mixed Discretion
2	- Mandatory Docket

states in which the judge is initially appointed to a lifetime term.¹⁷ In many states, these two terms are identical. These differences primarily occur in merit selection states, which

¹⁷The only major time this is an issue is in New Jersey, in which judges are initially appointed to a seven year term and then reappointed to a lifetime term. However, since New Jersey is represented with a seven year term initially, this is different than if they were initially appointed to a lifetime term and therefore their subsequent term is greater than ten years.

have one to three year initial terms followed by a retention election and then have longer terms after the first retention election. Coding term lengths as ordinal rather than continuous allows for consistent interpretation across the models. Retention Method ranges from zero to two and indicates the method of retention: zero indicates a simple reappointment procedure with no election. One represents an uncontested retention election, most common under the merit selection system, but also used in states that utilize partisan or non-partisan election systems. Two indicates a contested retention election, either partisan or non-partisan. Docket Control represents the amount of discretion that a court has over its docket. Zero is a completely discretionary docket, while two is no discretion. A one indicates a mix of discretionary and mandatory appeals. Table 2 shows the coding for all levels of each indicator.

3.2 Model

As [Linzer and Staton \(2014\)](#) show, a properly specified IRT model can be more informative than a simple additive index. The most important benefit of using this type of model is that there is no assumption made about the space between the cut points of each level of indicator, and between the indicators themselves ([Jackman, 2008](#); [Schnakenberg and Fariss, 2014](#)). This method also provides important tools in model specification and diagnostics that are lost when using an additive model. Many of the potential indicators of *de jure* judicial independence are dichotomous or nominal, having no ordering, and at best they are ordinal, as in the case of selection methods. Only the length of tenure is clearly ordered in that longer terms indicate independence as opposed to shorter terms. Other indicators such as selection, retention, and control of the docket, while theoretically grounded, do

not easily lend themselves to additive models. The details of the model used can be found in Appendix A.

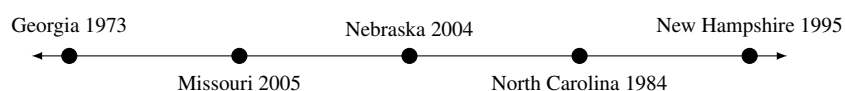
4 Results

I have specified two different models: a primary specification with five indicators and 42 years of data, and a regime change specification consisting of regime change observations over 200 years with four indicators.

Each model was run in R using Martyn Plummer’s JAGS and CODA packages for convergence diagnostics (R Core Team, 2014; Plummer et al., 2006; Plummer, 2014).¹⁸ Convergence is assessed using the Gelman-Rubin Diagnostic procedure using han Tsai, Gill and Rapkin’s *Superdiag* package in R (Gelman and Rubin, 1992; R Core Team, 2014).

Figure 7 shows how the results from this model map on to the Independence Continuum discussed in Section 1

Figure 7: Independence Continuum



4.1 Five Indicator Specification

The five indicator specification of this model is composed of all state year observations from 1970-2012. Figure 9 shows the results of an additive scale for the relevant indicators and data in this specification.

¹⁸Other software packages used are listed in the Reference section without being cited in-text.

Figure 8: *De Jure* Judicial Independence Over Time

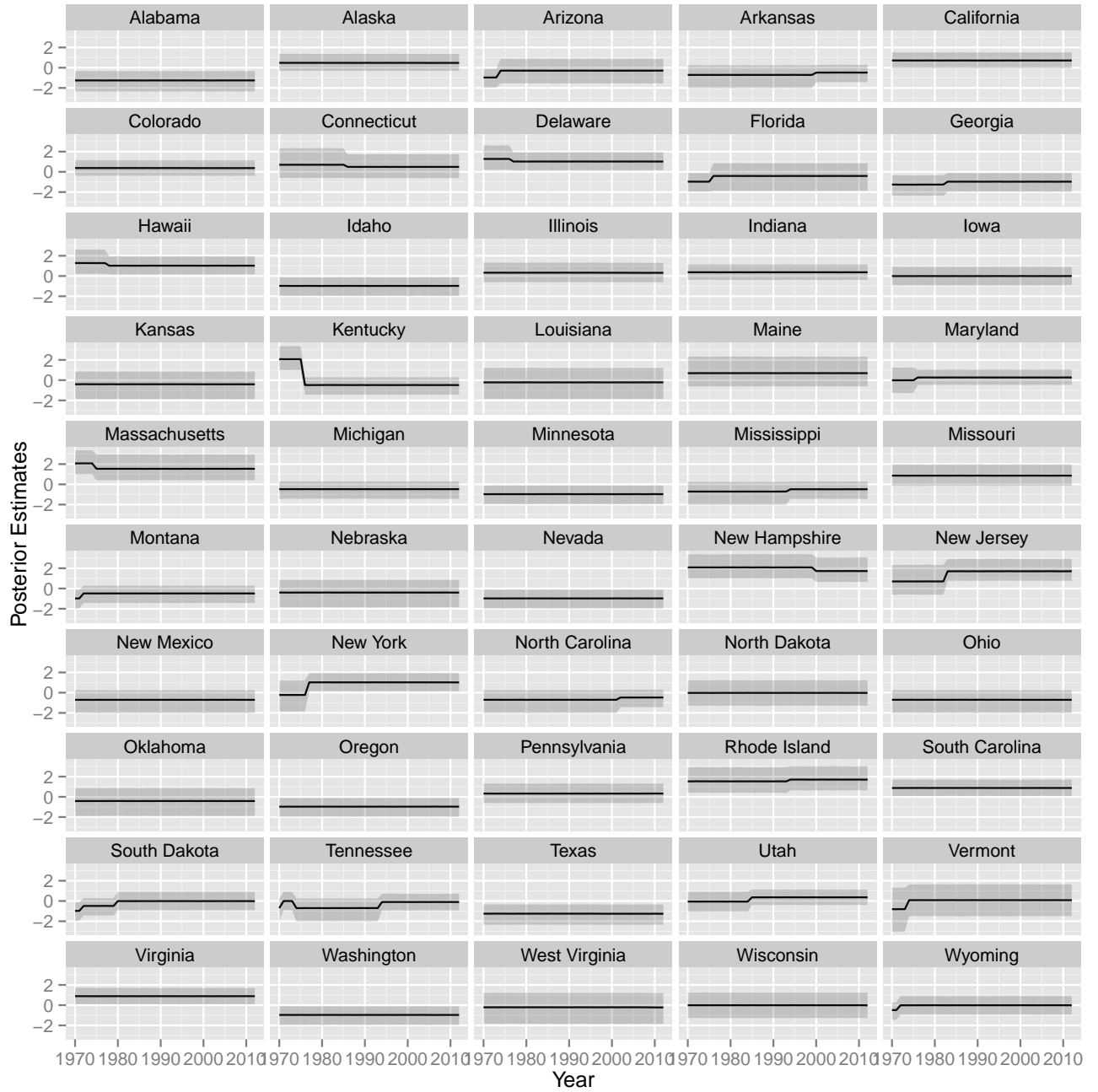


Figure 9: Five Indicator Specification- Additive vs. IRT

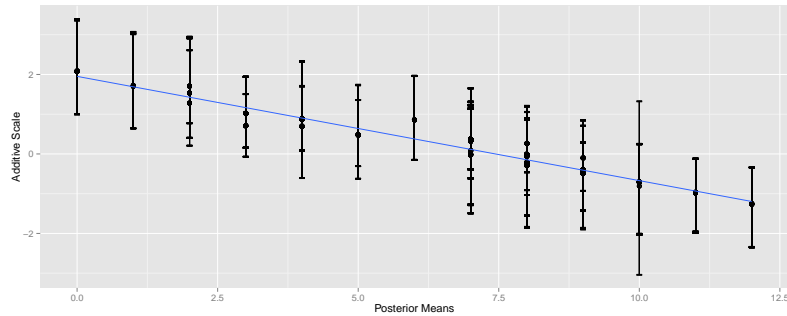
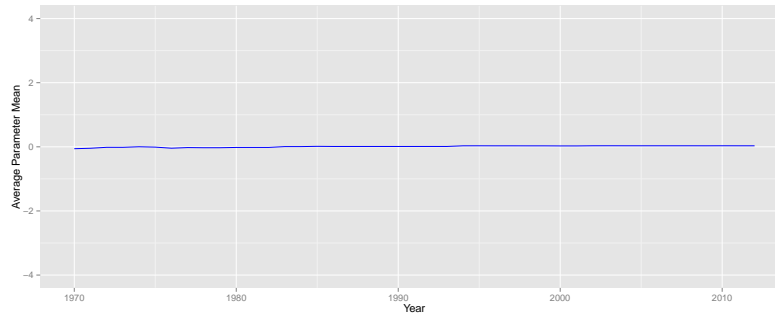


Figure 8 shows the results of the posterior estimates over the time period in the sample. Due to the short time period in the sample, there is little variance in the posterior estimates. However, there is variance between the indicators, which allows for the model to converge. The Regime Change Specification was run using 100,000 draws with the first 25,000 thrown away as burn-in. Convergence was assessed with a Gelman-Rubin Diagnostic, with a Multivariate Potential Scale Reduction Factor of 1.42 (Gelman and Rubin, 1992). Figure 10 shows the mean for the whole sample by year. Even during the short time period of this study, there has been a substantial change which has affected the sample mean. This indicates that since 1970, while some states have moved toward more independent systems, many more states have moved toward more accountable systems. The Pearson's Correlation between the additive and the IRT model is $-.9553$ with a p -value of $< .001$.

Figure 11 shows the Beta Discrimination for the indicators in the model. The beta discrimination is an evaluation of the affect on the slope of each indicator. The highest is for retention elections. This fits nicely into the theory of Choi, Gulati and Posner (2010) that control of retention elections are key to determining the independence and subsequent

Figure 10: Average Judicial Independence Over Time



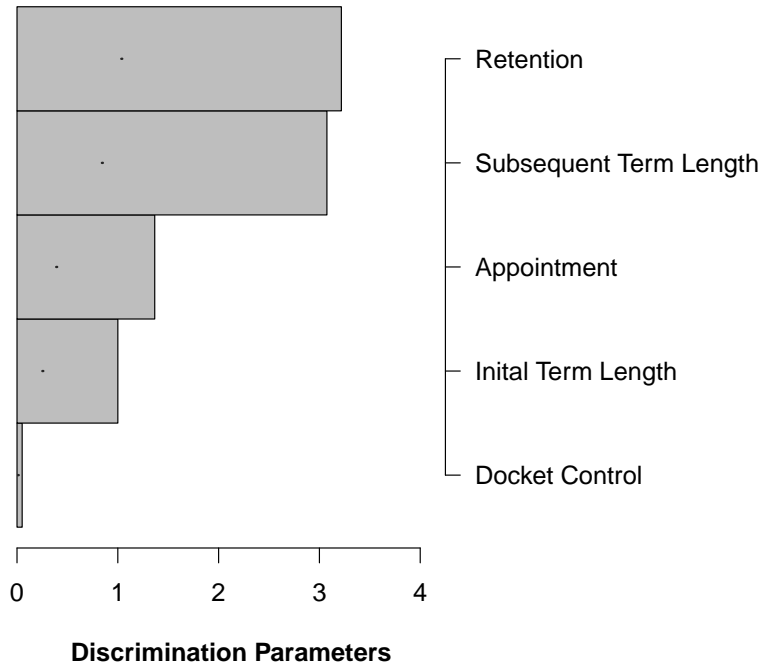
actions of a judge on the bench.

4.2 Regime Change Specification

The second model specification consists of the four original indicators, but is limited to only unique observations of each judicial institution rather than repeated observations over long periods of time. As shown in Figure 11, the omitted indicator, docket control, does not have a large effect on this model, so it does not suffer too much from this lack of data. This specification dramatically reduces the number of observations. For instance, Massachusetts, which has maintained the same selection/retention methods as well as the same term lengths for its entire history has only a single observation. There are 155 state-year observations in this model specification. The Regime Change specification was run using 10,000 draws with the first 5,000 thrown away as burn-in. Convergence was assessed with a Gelman-Rubin Diagnostic, with a Multivariate Potential Scale Reduction Factor of 1.04 ([Gelman and Rubin, 1992](#)).

Figure 13 shows the results of an additive scale for the relevant indicators in the Regime Change specification plotted against the IRT model. The Pearson's Correlation between

Figure 11: Beta Discrimination



the additive model and the IRT model is $-.899$, with a p -value of $\leq .001$.

5 Validation

5.1 Face Validity

As shown with the evidence above, this measurement model uses indicators that have been shown through previous literature to indicate judicial independence. In addition, these are all indicators that are institutional in nature, rather than measuring some concept that

Figure 12: Regime Change Specification- Parameter Means

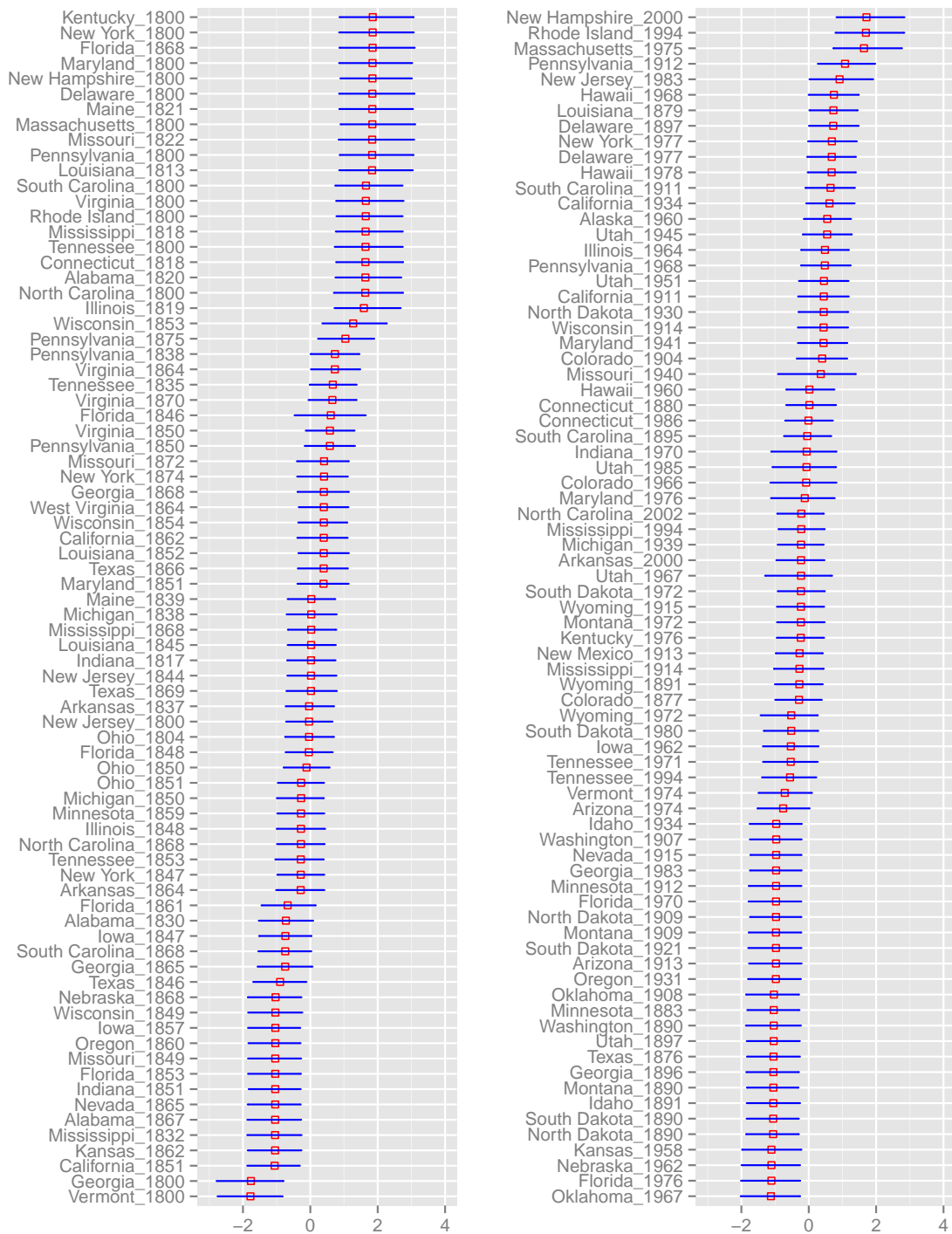
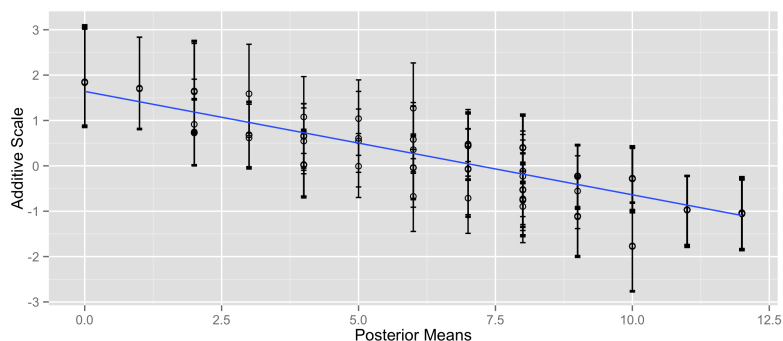


Figure 13: Regime Specification- Additive vs. IRT



would measure both *de facto* and *de jure* independence. Rather than measuring the amount of docket control that a court exercises, I examine the control that is constitutionally or statutorily authorized to the court.

5.2 Convergent Validity – Ginsburg and Melton Replication

As a test of convergent validity, I have replicated the additive measure discussed above from [Melton and Ginsburg \(2014\)](#). [Melton and Ginsburg \(2014\)](#) gathered data from the Comparative Constitutions Project (CCP) to code indicators of *de jure* judicial independence using the text of national constitutions. I replicated their project and applied it to the American states by using the texts of state constitutions gathered from the NBER/Maryland State Constitutions Project (SCP) ([Wallis, N.d.](#)).¹⁹ Of the many variables available in the CCP, [Melton and Ginsburg \(2014\)](#) used five questions to code an additive index of *de jure* judicial independence. The five questions consisted of: a statement of judicial independence, lifetime terms, selection procedures, removal procedures, removal conditions,

¹⁹The SCP did not have text for all current state constitutions, for the states that were unavailable, the official texts were downloaded from the state government website ([Wallis, N.d.](#)).

and salary insulation. The questions used by CCP are available in Appendix C. This data is coded using exclusively the text of the constitutions in question, as such, my replication has followed the same procedures.

Table 3: Correlation of Measures

	Additive	IRT	Melton and Ginsburg
Additive	–	–.95***	–0.45**
IRT	–0.95***	–	0.52***
Melton and Ginsburg	–0.45**	0.52***	–
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01			

The index created by Melton and Ginsburg (2014) was correlated with both the five indicator model and my additive index using Polychoric Correlation (Olsson, 1979).²⁰ The data used for these calculations are located in Appendix D. Table 3 shows these correlations. All are statistically significant, however, they are not overwhelmingly correlated.

Discriminant Validity

To test for any other factors that could potentially be associated with the indicators used for this measure of judicial independence I correlate both the IRT measure and the additive measure with Enns and Koch’s (2013) state-measure of policy mood. The results are shown in Table 4. Divergent validity is contrasted with convergent validity in that it ensures that the measure in question does not measure things that it is not supposed to measure (Campbell and Fiske, 1959; Jackman, 2008). As shown in Table 4 this criteria is satisfied to a respectable degree.

²⁰Correlation was also calculated using Pearson’s *R* and Kendall’s Tau *B* and were not significantly different.

Table 4: Correlation of Estimates with [Enns and Koch \(2013\)](#)

	Mood	2 Dimension Mood	Posterior Estimates	Additive
Mood	–	0.58***	0.40**	–0.40**
2 Dimension Mood	0.58***	–	0.16	–0.15
Posterior Estimates	0.40**	0.16	–	–0.95***
Additive	–0.40**	–0.15	–0.95***	–
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

6 Application

There is a variety of ways in which this measurement model can be applied in future research. For use outside of judicial politics, it provides a convenient means to control for judicial independence in the study of state politics. In a series of articles and books Gray and Lowery discuss the role of interest groups, lobbying, and money in state politics ([Berkhout et al., 2011](#); [Gray and Lowery, 2000](#); [Lowery and Gray, 1995](#)). However, none of these articles make mention of the courts. State courts are often integral players in campaign finance and lobbying reforms. Some of the most important factors in a reforms chances of being found constitutional is the individual judge's disposition to campaign donations. That disposition will be much more subject to change if the judge him or herself is forced to campaign, and is taking campaign donations from the same political action groups, groups, and lawyers that they are hearing cases against.

The influence of the judiciary has long been ignored in policy diffusion literature. Despite this, in their article on the state of policy diffusion research [Graham, Shipan and Volden](#) neglect to examine the role of the judiciary on the diffusion of policies throughout the state ([Graham, Shipan and Volden, 2013](#)). There has been a a considerable expansion of same-sex marriage legalization as well as bans on same sex marriage in the past decade.

Hundreds of judges, both federal and state, from district judges to Supreme Court justices have taken up this issue. An overwhelming number of states have found bans on same sex marriage constitutional while a wave of federal district and appellate judges after the Supreme Court's ruling in *United States v. Windsor* (133 S.Ct. 2675). What is the cause for the disparity between the state and federal judges? Certainly this isn't due to a misunderstanding of the Constitution between state and federal judges, or an entirely partisan issue.

7 Conclusion

Using a new dataset of constitutional features, I estimate a Bayesian Ordinal Item Response Theory model variable model measuring *de jure* judicial independence. I illustrate how these various institutional structures can affect a state supreme court's level of *de jure* judicial independence. I also illustrate how these institutions have changed over time, both within states and throughout the whole country.

While the short time period in the primary specification of the model would seem to indicate that judicial independence is a relatively static thing, the mean plot shows that it is significantly changing even over a short period of time.

There are two avenues of research that should be explored subsequent to this paper. The first is to establish a longer period of time in this model. This could be easily accomplished with the correct amount of computing time. The second extension of this research should be to expand this to a cross-national sample, which will allow for a greater variety of selection and retention methods.

References

- Adams, John. 1794. *A Defence of the Constitutions of Government of the United States of America: against the attack of M. Turgot in his Letter to Dr. Price, dated the twenty-second day of March, 1778*. Vol. 3 J. Stockdale.
- American Bar Association, Commission on the 21st Century Judiciary. 2003. "Justice in Jeopardy."
- Aspin, Larry. 2000. "Trends in Judicial Retention Elections, 1964-1998." *Judicature* 83:79–81.
- Berkhout, Joost, Brendan Carrol, David Lowery and Dmitr Toshkov. 2011. Timing is Everything? Organized Interests and the Timing of Legislative Activity. In *EUSA Twelfth Biennial International Conference, Boston, March*. pp. 3–5.
- Bonneau, Chris W and Melinda Gann Hall. 2010. *In Defense of Judicial Elections*. Routledge.
- Brace, Paul and Brent D. Boyea. 2008. "State Public Opinion, the Death Penalty, and the Practice of Electing Judges." *American Journal of Political Science* 52(2):360–372.
- Burbank, Stephen B. 2008. "Judicial Independence, Judicial Accountability and Inter-branch Relations." *Dadalus* 137(4):16–27.
- Caldarone, Richard P., Brandice Canes-Wrone and Tom S. Clark. 2009. "Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions." *The Journal of Politics* 71(2):560–573.
- Campbell, Donald T and Donald W Fiske. 1959. "Convergent and discriminant validation by the multitrait-multimethod matrix." *Psychological bulletin* 56(2):81.
- Canes-Wrone, Brandice, Tom S. Clark and Jee-Kwang Park. 2012. "Judicial Independence and Retention Elections." *Journal of Law, Economics, and Organization* 28(2):211–234.
- Carrington, Paul D. 1998. "Judicial Independence and Democratic Accountability in Highest State Courts." *Law and Contemporary Problems* 61(3):79–126.
- Choi, Stephen, G. Mitu Gulati and Eric A. Posner. 2010. "Professionals or Politicians: The Uncertain Empirical Case for an Elected Rather than Appointed Judiciary." *Journal of Law, Economics, & Organization* 26(2):290–336.

- Cingranelli, David L. and David L. Richards. 2008. "The Cingranelli-Richards (CIRI) Human Rights Dataset." Online.
URL: <http://www.humanrightsdata.org/>
- Debow, Michael, Diane Brey, Erick Kaardal, John Soroko, Frank Strickland and Michael B. Wallace. 2013. The Case for Partisan Judicial Elections. Technical report The Federalist Society.
- Eddelbuettel, Dirk. 2015. *random: True Random Numbers using random.org*. R package version 0.2.3.
URL: <http://CRAN.R-project.org/package=random>
- Enns, Peter K and Julianna Koch. 2013. "Public Opinion in the US States 1956 to 2010." *State Politics & Policy Quarterly* 13(3):349–372.
- Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108(02):297–318.
- 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* 19:497–527.
- Ferejohn, John and Pasquale Pasquino. 2003. "Constitutional adjudication: lessons from Europe." *Tex. L. Rev.* 82:1671.
- Fontana, David. 2011. Docket Control and Success of Constitutional Courts. In *Comparative Constitutional Law*, ed. Tom Ginsburg and Rosalind Dixon. Northampton, MA: Edward Elgar.
- Fox, John. 2010. *polycor: Polychoric and Polyserial Correlations*. R package version 0.7-8.
URL: <http://CRAN.R-project.org/package=polycor>
- Gelman, Andrew and Donald B Rubin. 1992. "Inference from iterative simulation using multiple sequences." *Statistical science* pp. 457–472.
- Gibson, James L. 2008. "Challenges to the Impartiality of State Supreme Courts: Legitimacy Theory and "New-Style" Judicial Campaigns." *American Political Science Review* 102(1):59–75.

- Graham, Erin R, Charles R Shipan and Craig Volden. 2013. "The diffusion of policy diffusion research in political science." *British Journal of Political Science* 43(03):673–701.
- Gray, Virginia and David Lowery. 2000. *The population ecology of interest representation: Lobbying communities in the American states*. University of Michigan Press.
- Hall, Melinda. 2014. *Attacking Judges: How Campaign Advertising Influences State Supreme Court Elections*. Stanford University Press.
- Hall, Melinda Gann. 1985. "Docket Control as an Influence on Judicial Voting." *Justice System Journal* 10(2):243–255.
- Hall, Melinda Gann. 1987a. "Constituent influence in state supreme courts: Conceptual notes and a case study." *The Journal of Politics* 49(04):1117–1124.
- Hall, Melinda Gann. 1987b. "An Examination of Voting Behavior in the Louisiana Supreme Court." *Judicature* 71(1):40–46.
- Hall, Melinda Gann. 2007. Competition as Accountability in State Supreme Court Elections. In *Running for Judge: The Rising Political, Financial, and Legal Stakes of Judicial Elections*, ed. Matthew Streb. New York: New York University Press.
- han Tsai, Tsung, Jeff Gill and Jonathan Rapkin. 2012. *superdiag: R Code for Testing Markov Chain Nonconvergence*. R package version 1.1.
URL: <http://CRAN.R-project.org/package=superdiag>
- Hayo, Bernd and Stefan Voigt. 2007. "Explaining de facto Judicial Independence." *International Review of Law and Economics* 27:269–290.
- Howard, Robert M. and Henry F. Carey. 2004. "Is an Independent Judiciary Necessary for Democracy?" *Judicature* 87(6):284–292.
- Jackman, Simon. 2008. Measurement. In *The Oxford handbook of political methodology*, ed. Janet M Box-Steffensmeier, Henry E Brady and David Collier. Oxford Handbooks Online chapter Measurement, pp. 1–59.
- Judicial Selection in the States*. 2014. Online.
URL: <http://www.judicialselection.us/>
- Keith, Linda Camp. 2002a. "Constitutional Provisions for Individual Human Rights (1977-1996): Are They More than Mere "Window Dressing?"." *Political Research Quarterly* 55(1):111–143.

- Keith, Linda Camp. 2002*b*. "Judicial Independence and Human Rights Protection Around the World." *Judicature* 85(4):195–200.
- La Porta, Rafael, Florencio Lopez de Silanes, Cristian Pop-Eleches and Andrei Shleifer. 2004. "Judicial Checks and Balances." *Journal of Political Economy* 112:445–70.
- Lang, Duncan Temple. 2007. "R as a Web Client—the RCurl package." *Journal of Statistical Software*, <http://www.jstatsoft.org> .
- Linzer, Drew A. and Jeffrey K Staton. 2014. "A Measurement Model for Synthesizing Multiple Comparative Indicators: The Case of Judicial Independence."
- Lowery, David and Virginia Gray. 1995. "The population ecology of Gucci Gulch, or the natural regulation of interest group numbers in the American States." *American Journal of Political Science* pp. 1–29.
- Maltzman, Forrest, James F Spriggs and Paul J Wahlbeck. 2000. *Crafting Law on the Supreme Court: The Collegial Game*. Cambridge University Press.
- Marshall, Monty G and Keith Jagers. 2002. "Polity IV project: Political regime characteristics and transitions, 1800-2002."
- Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.
- May, Melissa S. 2013. "Judicial Retention Elections After 2010." *Ind. L. Rev.* 46:59.
- McNollgast. 2006. "Conditions for Judicial Independence." *Journal of Contemporary Legal Issues* 105.
- Melton, James and Tom Ginsburg. 2014. "Does De Jure Judicial Independence Really Matter?: A Reevaluation of Explanations for Judicial Independence." *Journal of Law and Courts* 2(2):187–217.
- Nelson, Michael J., Rachel Paine Caufield and Andrew D. Martin. 2013. "OH, MI: A Note on Empirical Examinations of Judicial Elections." *State Politics & Policy Quarterly* 13(4):495–511.
- Olson, Mancur. 2009. *The logic of collective action*. Vol. 124 Harvard University Press.
- Olsson, Ulf. 1979. "Maximum likelihood estimation of the polychoric correlation coefficient." *Psychometrika* 44(4):443–460.

- Plummer, Martyn. 2014. *rjags: Bayesian graphical models using MCMC*. R package version 3-14.
URL: <http://CRAN.R-project.org/package=rjags>
- Plummer, Martyn, Nicky Best, Kate Cowles and Karen Vines. 2006. "CODA: Convergence Diagnosis and Output Analysis for MCMC." *R News* 6(1):7–11.
URL: <http://CRAN.R-project.org/doc/Rnews/>
- R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: <http://www.R-project.org/>
- R Core Team. 2015. *foreign: Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, ...* R package version 0.8-63.
URL: <http://CRAN.R-project.org/package=foreign>
- Reid, Tracial V. 1999. "The Politicization of Retention Elections: Lessons from the Defeat of Justices Lanphier and White." *Judicature* 83:68–77.
- Ríos-Figueroa, Julio. 2006. *Judicial independence: Definition, measurement, and its effects on corruption. An analysis of Latin America*. ProQuest.
- Ríos-Figueroa, Julio and Jeffrey K Staton. 2014. "An Evaluation of Cross-National Measures of Judicial Independence." *Journal of Law, Economics, and Organization* .
- Rosenberg, Gerald N. 1991. *The Hollow Hope: Can Courts Bring About Social Change?* Chicago: The University of Chicago Press.
- Rottman, David B., Carol R. Flango, Melissa T. Cantrell, Randall Hansen and Neil LaFountain. 1998. State Court Organization 1998. Technical report Bureau of Justice Statistics Washington, D.C.: .
- Rottman, David B., Carol R. Flango and R. Shedina Lockley. 1993. State Court Organization 1993. Technical report Bureau of Justice Statistics Washington, D.C.: .
- Rottman, David B. and Shauna M. Strickland. 2004. State Court Organization 2004. Technical report Bureau of Justice Statistics Washington, D.C.: .
- Russell, Peter H and David M O'Brien. 2001. *Judicial independence in the age of democracy: Critical perspectives from around the world*. University of Virginia Press.

- Schmidhauser, John R. 1987. *Comparative Judicial Systems: Challenging Frontiers in Conceptual and Empirical Analysis*. Butterworth-Heinemann chapter Alternative conceptual frameworks in comparative cross-national legal and judicial research, pp. 34–57.
- Schnakenberg, Keith E and Christopher J Fariss. 2014. “Dynamic Patterns of Human Rights Practices.” *Political Science Research and Methods* 2(01):1–31.
- Schwab, Klaus and Xavier Sala-i-Martin. 2014. The Global Competitiveness Report 2014–2015. Technical report World Economic Forum.
- Songer, Donald R., Martha Humphries Ginn and Tammy A. Sarver. 2003. “Do Judges Follow the Law When There is No Fear of Reversal?” *Justice System Journal* 24(2):137–161.
- Squire, Peverill. 2008. “Measuring the Professionalization of U.S. State Courts of Last Resort.” *State Politics & Policy Quarterly* 8(223–238).
- Su, Yu-Sung and Masanao Yajima. 2015. *R2jags: A Package for Running 'JAGS' from R*. R package version 0.05-03.
URL: <http://CRAN.R-project.org/package=R2jags>
- Tiede, Lydia Brashear. 2006. “Judicial Independence: Often Cited, Rarely Understood.” *Journal of Contemporary Legal Studies* 15:129–161.
- Treier, Shawn and Simon Jackman. 2008. “Democracy as a Latent Variable.” *American Journal of Political Science* 52(1):201–217.
- United States v. Windsor*. 133 S.Ct. 2675.
- Varnum v. Brien*. 763 N.W.2d 862 Iowa 2009.
- Voeten, Erik. 2008. “The impartiality of international judges: Evidence from the European Court of Human Rights.” *American Political Science Review* 102(04):417–433.
- Wallis, John Joseph. N.d. “The NBER/Maryland State Constitution Project.”
URL: <http://www.stateconstitutions.umd.edu/index.aspx>
- Watson, Richard A and Rondal Gene Downing. 1969. *The Politics of the Bench and the Bar: Judicial Selection Under the Missouri Nonpartisan Court Plan*. John Wiley & Sons.

Wickham, Hadley and Romain Francois. 2015. *dplyr: A Grammar of Data Manipulation*. R package version 0.4.1.
URL: <http://CRAN.R-project.org/package=dplyr>

A Model Specification

The model used for this is a Bayesian Dynamic-Ordinal Item Response Theory Model. This model is the same as that developed by [Schnakenberg and Fariss \(2014\)](#). Rather than examining countries' respect for human rights, this model examines the level of *de jure* judicial independence that a state's judicial institutions represents.²¹

I assume that the observed indicators for each state-year are functions of a single dimensional latent variable (the Judicial Independence Continuum) that represents the level of *de jure* judicial independence. For each state-year observation, let i index the state and t index the year. For each model, there are J indicators $J = 1, \dots, J$ each of which is ordinal.²² I estimate each θ_{it} , which is the latent level of *de jure* judicial independence of each state i in year t ([Schnakenberg and Fariss, 2014, 7](#)).

Let $i = 1, \dots, N$ index cross-sectional units and $t = 1, \dots, T$ index time periods. In each time, period, I observe values y_{ij} for each of $j = 1, \dots, J$ indicators for each unit. Each indicator is assumed to be ordinal and can take on K_j values.²³ The responses to each of the items depend on a single latent variable θ_{it} , which may vary across units and over time. I assume that all indicators are independently drawn from a logistic distribution ([Schnakenberg and Fariss, 2014, 7](#)).

For each indicator, there is an item discrimination parameter β_j and a set of $K_j - 1$ difficulty cut-points ([Schnakenberg and Fariss, 2014, 7](#)).²⁴ A benefit of using institutional arrangements, which are observable with readily available data is that any error that is introduced into the indicators is the result of coding errors rather than perceptual errors found in survey responses.

²¹There is also one major difference in that [Schnakenberg and Fariss](#) measure a *de facto* level of respect for human rights, in that they look at actual numbers of torture, extra-judicial killings, and disappearances in a given year, I am looking at institutional arrangements, which typically have much less within-state variance from year to year.

²²For definitions of each level of indicators, see Table 2

²³For term length indicators, these have been collapsed into five ordered categories. This is due to the relative infinite value of a life or quasi-life term. A previous specification of this model was considered using the average length of term in office for each justice that had served on the court under the life or quasi-life term. This specification was rejected as introducing a questionable degree of *de facto* independence into what is a purely *de jure* measurement model.

²⁴For more detail on these cut-points see [Treier and Jackman \(2008\)](#); [Schnakenberg and Fariss \(2014\)](#).

The probability distribution for a given response to item j is observed as:

$$P[y_{ij} = k] = F(\alpha_{jk} - \theta_{it}\beta_j) - F(\alpha_{jk-1} - \theta_{it}\beta_j) \quad (1)$$

Where $F(\cdot)$ is the logistic cumulative distribution function. Assuming the local independence of responses across units, the likelihood function for β , α , and θ given the data is:²⁵

$$\mathcal{L}(\beta, \alpha, \theta|y) = \prod_{i=1}^N \prod_{t=1}^T \prod_{j=1}^J [F(\alpha_{jy_{it}} - \theta_{it}\beta_j) - F(\alpha_{jy_{it}-1} - \theta_{it}\beta_j)] \quad (2)$$

If θ , (*de jure* judicial independence), was able to be observed, the likelihood function shown in Equation 2 would be equivalent to an independent ordinal logistic regression model (Schnakenberg and Fariss, 2014, 8).

There are three basic independence assumptions considered for this IRT model. The assumptions are: (1) local independence of different indicators within the same state – year, (2) local independence of indicators across states within years and (3) local independence of indicators across years within states (Schnakenberg and Fariss, 2014, 8). The implication of this is that each y_{itj} are independently drawn conditional on θ . The only relationship between two item responses is that they are both measure the same latent variable, θ . Schnakenberg and Fariss (2014, 8) discuss the assumptions made in this model in greater detail. I relax Assumption 1 slightly in that there is some correlation between selection and retention methods. For example, there are no states which use a commission system for initial selection and then require judges to run in competitive elections.²⁶ However, their Assumption 3, is a matter of concern in this model for the same reason as it was in Schnakenberg and Fariss (2014). Assumption 3 relates to the local independence of indicators across years within states (Schnakenberg and Fariss, 2014, 8). Rather than attempting to measure a regime’s respect for human rights as in Schnakenberg and Fariss (2014), I am measuring an institutional arrangement by each state. In each state there is a yearly potential to change selection methods, keeping intact a relaxed version of this assumption. This relaxation of Assumption 3 comes from the informative priors used by

²⁵For a more detailed explanation of the assumptions made in this likelihood, see Schnakenberg and Fariss (2014, 8).

²⁶This assumption does not extend to filling judicial vacancies. For example, New Mexico uses a merit selection system to fill judicial vacancies and then requires the incumbent to run in a competitive partisan election at the next general election (*Judicial Selection in the States*, 2014). I consider this to be an artifact of the *de facto* nature rather a factor of *de jure* judicial institutions.

B Coding Notes

- Arkansas- In 1868, the Chief Justice was appointed by the Governor, with Senate Consent, with the Associate Justices elected by the people. This is coded as partisan election since the majority of the justices are elected.
- Connecticut- from 1784 through 1818 Judges were appointed by the Legislature, but the tenure in office was dependent upon the action of the legislature. These years are coded as missing and subsequently dropped from the dataset.
- Delaware- The original court of last resort was the Court of Appeals, and the data is coded as such. The Supreme Court was created by constitutional amendment in 1951, at which point the data is coded to reflect the Supreme Court.
- Michigan- In 1939, a constitutional amendment passed calling for non-partisan elections for judges, except the Supreme Court, which would continue to be nominated at party conventions. This is coded as a non-partisan following practice in this field.
- New Jersey- AJS Data does not reflect the Constitutional amendment in 1983, which extended subsequent terms to a term of good behavior.
- New Mexico- Judges are elected for the remainder of the unexpired term. This could be up to 8 years, so it was coded as such.
- Oklahoma is examined individually, but both courts have mandatory jurisdiction.- 2004
- Pennsylvania- From 1874 to 1968 Supreme Court justices were elected to twenty-one year terms and were not eligible for reelection. This is coded as a 0 for the subsequent term.
- Tennessee- Constitution says qualified electors, however, by executive order the governor has created a nominating commission
- Texas is examined individually, SC has discretionary jurisdiction, but COA has mandatory in sentencing issues, so this is coded as mixed.- 2004
- Retention Elections- Gubernatorial reappointment, judicial commission reappointment, legislative reappointment are coded as No Retention Elections.

- Docket Control- This is established by looking at Criminal and Civil cases. Administrative agency decisions are omitted, as well as death penalty cases.
- Docket Control- Observations between BJS reports are assumed to be static.

C Comparative Constitutions Project Variables

- Statement of Judicial Independence – **[JUDIND]** – Does the constitution contain an explicit declaration regarding the independence of the central judicial organ(s)?
- Life Term – **[SUPTERM]** – What is the maximum term length for judges for the highest ordinary court?
- Selection Procedure – **[SUPNOM,SUPAP]** – Who is involved in the nomination of judges to the highest ordinary court? and Who is involved in the approval of nominations to the highest ordinary court?
- Removal Procedures – **[JREM, JREMPRO, JREMFIRP, JREMSECP, JREMBOTP, JREMAP]** – Are there provisions for dismissing judges? and Who can propose the dismissal of judges? and Who can approve the dismissal of judges?
- Removal Conditions – **[JREMCON]** – Under what conditions can judges be dismissed?
- Salary Insulation – **JUDSAL** – Does the constitution explicitly state that judicial salaries are protected from governmental intervention?

D Ginsburg and Melton Replication Data

State	Additive	IRT	G&M	State	Additive	IRT	G&M
Alabama	12	-1.31091129	1	Montana	9	-0.581144339	1
Alaska	5	0.368783882	4	Nebraska	9	-0.072441168	2
Arizona	8	-0.012413602	4	Nevada	11	-0.909487793	3
Arkansas	9	-0.578075634	3	New Hampshire	1	1.533583382	2
California	3	0.673908309	4	New Jersey	2	1.768516335	3
Colorado	7	0.384204611	3	New Mexico	10	-0.905869398	1
Connecticut	5	0.760455328	3	New York	3	1.103325872	3
Delaware	3	1.099737375	3	North Carolina	9	-0.576762507	3
Florida	9	-0.076706106	3	North Dakota	7	-0.291201696	3
Georgia	11	-0.908038872	2	Ohio	10	-0.89905174	3
Hawaii	3	1.100137397	4	Oklahoma	9	-0.079603662	3
Idaho	11	-0.913301035	1	Oregon	11	-0.903450105	1
Illinois	7	0.16588657	3	Pennsylvania	7	0.169291474	3
Indiana	7	0.387192126	4	Rhode Island	1	1.520687113	5
Iowa	8	0.155611872	4	South Carolina	4	0.900738263	4
Kansas	9	-0.079342362	4	South Dakota	8	0.145693754	3
Kentucky	9	-0.58387198	1	Tennessee	9	0.022875851	3
Louisiana	8	-0.572016606	3	Texas	12	-1.309764875	2
Maine	4	1.056674084	4	Utah	7	0.377535504	4
Maryland	8	0.241124912	4	Vermont	7	0.521552848	3
Massachusetts	2	1.30252114	2	Virginia	4	0.909315298	3
Michigan	9	-0.579248113	3	Washington	11	-0.900776219	2
Minnesota	11	-0.90983707	1	West Virginia	8	-0.576826557	3
Mississippi	9	-0.585931897	3	Wisconsin	7	-0.281544702	0
Missouri	6	0.663536048	3	Wyoming	8	0.152640428	4

E Correlation for Indicators

Table 5: Correlation of Indicators

	APPT	TERM1	TERM2	RETELE
APPT	—	0.28***	0.41***	0.81***
TERM1	0.28***	—	0.86***	0.42***
TERM2	0.41***	0.86***	—	0.49***
RETELE	0.81***	0.42***	0.49***	—
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			