Journal of Law, Economics, and Organization Advance Access published March 30, 2015

Trial and Error: Decision Reversal and Panel Size in State Courts*

Yosh Halberstam
University of Toronto

Using cross-state and within-court variation, I show that lower court decisions are reversed more frequently by larger, rather than smaller, panels of high court judges. Overall, conditional on being reviewed, the probability that a case is reversed by a high court judicial panel is less than one half. To understand these findings, I develop a simple framework that connects reversals and panel size with the extent to which judicial decision-making is congruent with the law. Assuming the high court rules correctly more often than not, my empirical results suggest that increasing judicial panel size erodes the quality of decision-making in high courts. These results are consistent with a large literature investigating small group size effects on productivity and output. (*JEL* D02, D71, H41, K40)

1. Introduction

Individual contributions to public goods depend on group size. In particular, the efficacy of small decision-making bodies, such as corporate boards and juries, appears to crucially depend on their size. For example, research in finance has shown that productivity and firm valuation is lower among firms with larger boards of directors. Likewise, studies in social psychology and management have found adverse size-effects on the psychology, communication, and organizational behavior of teams. In this article, I ask whether judicial panel size similarly affects outcomes.

In the United States, the number of judges who sit to hear a case on the highest state court varies.[1] Although appeals to the highest courts are typically reviewed by the entire membership of the court (*en banc*), some

*I thank the co-editor, Jennifer Reinganum, and two anonymous referees for their very helpful feedback. David Austen-Smith, Ernesto Dal Bo, Tim Feddersen, Tom Hubbard, Brian Knight, Kory Kroft, Jeff Mandell, Bob Powell, Bill Rogerson, Emily Satterthwaite, Michael Smart, Colin Stewart, Matt Turner, and participants at the University of Chicago (Harris), Tel Aviv University and Canadian Economic Association, and Canadian Law and Economics Association meetings provided helpful comments and suggestions. I am grateful to numerous state court clerks and officials for their input, as well as Graham Beattie and Adriana Robertson who provided outstanding research assistance.

1. These courts usually, but not always, are called the state supreme court. Other descriptions for the highest court within a legal jurisdiction include court of last resort and highest court of appeal. In general, the decisions of such a court are not subject to review by another court.

The Journal of Law, Economics, and Organization, Vol. 0, No. 0
doi:10.1093/jleo/ewv007

states have five high court judges, some states seven, and the remaining states have nine high court judges. This is in contrast to lower state courts, where one judge hears a case at the trial court and, in most states, three judges hear a case at the intermediate appellate court.

Using data on US state high court decisions, I find that reversal rates of lower court decisions are (a) uniformly less than one half and (b) increasing in the number of judges on the reviewing panel. To understand how these findings connect to welfare, I develop a simple framework with two states, one in which the law favors affirmation and one in which it favors reversal. The key identifying assumption of the model is that panel size is randomly assigned to cases. Given this, the theory then implies that higher reversal rates correspond to worse decision-making by the reviewing judicial panels.

I use three sources of variation to corroborate my findings. First, using approximately 5000 decisions made by sub-panels of high court judges (within-court), I find that a marginal increase of one high court judge increases the probability that the lower court decision being reviewed is reversed by 2 percentage points. Second, regression analysis of case-level data on over 20,000 state supreme court *en banc* decisions between 1995 and 1998 indicates that, relative to a panel of five judges, which reverses 28% of lower court decisions, a marginal panel increase of two judges is associated with 3–4 percentage points greater likelihood of reversal. Likewise, a panel of nine judges is 9–10 percentage points more likely to reverse a lower court's decision than a panel of seven judges. Third, I further corroborate these results using state-level annual data on reversals by state high courts between 2000 and 2011. I evaluate the robustness of my empirical results by including a variety of controls for type of case considered, panel demographics, judicial selection controls, and fixed effects. I find that the relationship between panel size and reversals remains robust.

To further explore the possibility that appeal selection might be driving the results, I examine three leading threats to identification in the cross-court regressions: that larger high courts are better positioned to detect erroneous decisions during the appeal selection process because they (a) have more resources, (b) have a more stringent appeal selection method, or (c) are higher quality. To test these theories, I collected data from multiple sources and surveyed state court clerks and other officials. However, I find only marginal support for these mechanisms and the main effects of panel size on reversal rates remain predominantly unaffected.

In the next section, I provide a summary of the relevant literature followed by the theoretical framework and institutional background on state courts. Section 5 describes the data I use to test the relationship between panel size and reversals. In Section 6, I present my empirical approach and, in Section 7, the results. In Section 8, I explore appeal selection mechanisms followed by robustness checks. Section 10 concludes.

2. Literature

Decision reversals are of substantial interest among legal scholars and practitioners. Legal scholars argue that decision reversals (and the threat of being reversed) are a primary driver in judges' decision-making calculus. Often, information on whether lower court decisions are affirmed (or reversed) is used to evaluate a court's quality and may ultimately influence its judges' path to promotion. Even if promotion is not on the horizon, judges do not like to be reversed (Posner 2005) and the desire to avoid being reversed by higher courts may significantly affect the opinions lower court judges write and the verdicts they reach (Watson 1987; Drahozal 1998).[2] Relative to this literature, I apply data on reversals to measure the quality of decision-making of the reviewing high court judges, as opposed to the judges reviewed by them.

Turning to the literature in law and economics, papers on appeal selection primarily model the interaction between different court levels. For example, Shavell (1995) and Daughety and Reinganum (2000) look at strategic aspects of appeal selection, treating the appeal process as an opportunity to correct erroneous decisions by lower court judges. In contrast, appeals can be viewed as a disciplinary tool to preserve ideological alignment within a legal hierarchy (Cameron et al. 2000). Others, such as Daughety and Reinganum (2006) and Spitzer and Talley (2013), explore within-panel strategic behavior that results from preference heterogeneity among appellate judges. At the same time, the literature on judicial panel size is sparse: Kornhauser and Sager (1986) discuss tradeoffs in panel size; George and Guthrie (2008) suggest that a panel of three high court judges is optimal. Relatedly, Alarie et al. (2011) look at optimal panel size in the context of the Supreme Court of Canada, where the chief justice selects the size and allocation of judges to panels. However, I am unfamiliar with work that looks at the role of group size in judicial outcomes.

More broadly, the effect of panel size on collective decisions and its welfare implications have been studied in various settings. A body of research in finance is dedicated to the study of corporate governance, focusing specifically on the relationship between the value of the firm and the size of its board of directors. For example, a seminal paper by Yermack (1996) finds that firms with smaller boards of directors have higher market valuation. Similar links have been documented in related studies. For example, Eisenberg et al. (1998) and Coles et al. (2008) extend the analysis and obtain similar findings in small firms and firms with simple organizational structures, respectively, and Mak and Kusnadi (2005) replicate these results in non-US firms. Likewise,

2. For example, Choi et al. (2012) find that district court judges choose to publish less yet higher-quality opinions when facing less cohesive (i.e., more unpredictable) circuit appellate panels. This is because published opinions are made public in which case their reversal carries a greater penalty to reputation relative to unpublished opinions. They also suggest that the overall effect of unpredictability is a higher reversal rate.

Huther (1997) finds that variable costs in firms, a measure of efficiency, decline in board size. Unlike firm valuation, the social value generated by the courts is unobservable. I contribute by developing a framework in which reversals can be applied to infer the relationship between the value of judicial panels and size.

Relatedly, the empirical literature on judicial decision-making has focused primarily on measuring the impact of judicial selection systems on, for example, the severity of rulings in criminal cases (Huber and Gordon 2004, 2007; Hall 1992; Lim 2013) and the quality of legal reasoning in opinions (Hall and Bonneau 2006; Choi et al. 2010). Much of the struggle in this literature focuses on constructing the appropriate observable and objective measure of judicial performance (see, e.g., Choi and Gulati 2006; Choi et al. 2008). More recent empirical work has focused on how judicial selection interacts with heterogenous preferences and private information on the courts (Iaryczower and Shum 2012; Iaryczower et al. 2013). In contrast, this article abstracts from judicial selection effects and explores the connection between judicial group size and outcomes.

A large literature on committee design highlights the pitfalls of adverse selection and moral hazard. For example, Palfrey and Rosenthal (1984) show that if agents are strategic, then free riding (moral hazard) can be pervasive. Likewise, Austen-Smith and Banks (1996) point out the possible shortcomings of the Condorcet Jury Theorem (Condorcet 1785) in the presence of heterogenous preferences (adverse selection). Related theoretical literature investigates the effects of free riding when information about a state-relevant payoff is costly to obtain, with some papers focusing on the optimal voting rule (Persico 2004; Gerardi and Yariv 2008; Gershkov and Szentes 2009) and others on the optimal number of decision-makers (Mukhopadhaya 2003; Martinelli 2006; Koriyama and Szentes 2009). However, there is little theory on optimal design in hierarchical institutions, such as courts. This article offers a theory that begins to address the mechanisms underlying such settings.

Finally, research in social psychology and management explores the behavioral effects of group size on individual and group behavior, such as conflict management, social perception, team productivity, team building, group decision-making, mediation, and group cohesiveness. Survey data and experimental evidence suggest that group size adversely affects outcomes. For example, Wheelan (2009) finds a negative relationship between group size and development, cohesion, and productivity of work groups, both in for profit and nonprofit firms. Likewise, Alnuaimi et al. (2010) find that team size has a negative effect on student performance, and that subjects placed in larger teams attributed more blame to one another, and took less responsibility for outcomes.

3. Background on State Courts

Most states have a three-tier hierarchy of courts. The lowest level of trial courts is tasked with a fact finding mission. *Trial courts* establish the facts in each case and, based on the expertise of the judge, apply the law to these facts to determine an outcome. The *intermediate appellate court* in each state discerns whether the law was applied appropriately in the trial. Typically, an appeal to the intermediate court is a matter of right. At the highest level in the hierarchy, the *state supreme court* is the court of last resort, which usually exercises discretion over which appeals to review.

The key distinction between the trial courts and appellate courts is that the trial court establishes the facts of each case. These facts are taken as given by the appellate courts upon review of an appeal. However, the appellate courts do not defer to the trial court's interpretation of the law (Posner 1998). Specifically, a state's high court evaluates whether the law was applied correctly in the trial and based on its evaluation will either affirm or reverse the lower court's decision (generally the intermediate appellate court unless there is none or if direct appeal to the highest court is possible).

The number of judges who sit to hear a given case on each court varies. In the context of comparative court systems, the number of judges on a panel often increases with the level of the court, particularly in hierarchical systems in which the decisions of judges in lower courts can be reviewed by a higher court upon appeal.[3] In the United States, panel sizes of intermediate appellate courts typically comprise three judges per case reviewed; whereas, trial courts, the lowest courts in the hierarchy, assign one judge to each case.[4] In contrast, state supreme courts comprise five to nine judges, where appeals are typically reviewed by the entire membership of the court (*en banc*).

As preliminary evidence on the relationship between reversals and panel size, in Figure 1, I plot the proportion of cases that were reversed by state supreme courts in each of the 4 years from 1995 to 1998 by panel size. In the 50 states, 17 high courts are comprised of five judges, 28 are comprised of seven judges, and 5 are comprised of nine judges. The pattern that emerges is unambiguous: the rate of reversal appears strongly (positively) correlated with panel size and is uniformly less than one half.

3. The number of judges who sit on a panel to hear a legal matter varies, often by the type of case being heard and the level of the court. In some courts, such as the Supreme Court of Canada, the number of judges assigned to a panel is determined on a case-by-case basis. In other systems, such as the federal courts of the United States, such variance is rarely present.

4. Appeals may be heard by more than three judges in 3 of the 40 states with intermediate appellate courts.

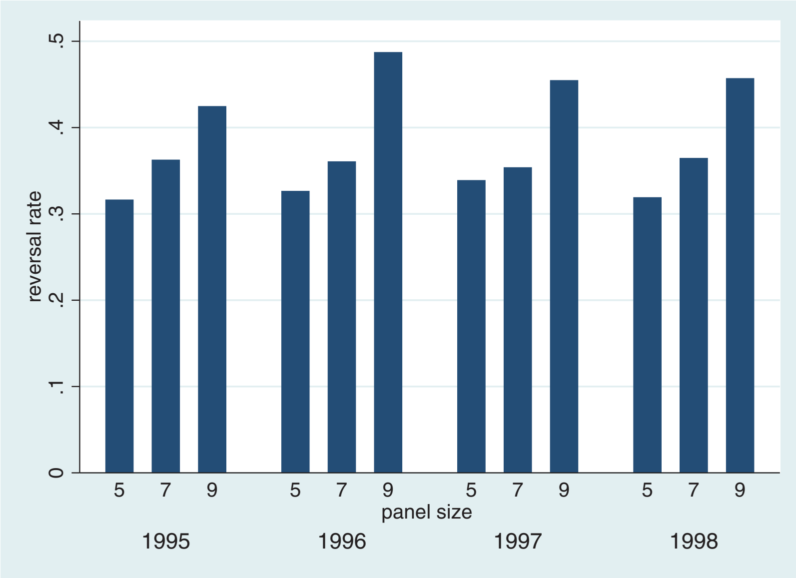

Figure 1. Reversals of Lower Court Rulings.
Notes: This figure plots the raw means of reversal rates per year by state supreme court panel size using case-level data from the State Supreme Court Data Project. The horizontal axis displays the number of state high court justices; the vertical axis denotes the proportion of lower court decisions reversed.

4. Theoretical Framework

Consider a panel of n high court judges that can either reverse or affirm lower court decisions. I assume that whether a case warrants a reversal or not is independent of panel size. Specifically, let $1-q$ denote the probability that the law favors a decision reversal by the panel.[5] Given that reversal rates are uniformly less then one half, let $\rho_n < 1/2$ denote the rate at which a panel of size n reverses lower court decisions. Thus, if ρ_n diverges from $1-q$, then outcomes generated by the panel are socially suboptimal.

Turning to beliefs about courts, it seems natural to assume that judicial panels are more likely to rule correctly than incorrectly on any given case (i.e., reverse the lower court when the law favors reversal and affirm otherwise). Denote this probability by $\phi_n > 1/2$. Given this, we have the

5. In subsequent sections, I address the validity of this assumption empirically. The greatest threat to identification is that q_n increases in n due to any number of institutional or behavioral reasons. To address this threat, I show that my results are robust to using within-court variation, where the assignments of appeals to panels are as good as random. I also focus on appeal selection (q_n) models that suggest how panel size might be correlated with the distribution of dockets that end up in the set of appeals pending review by the high court.

following relationship between reversals and the probability that a panel of n judges rules correctly:

$$\rho_n = \underbrace{q(1-\phi_n)}_{\text{reversal of correct decision in trial}} + \underbrace{(1-q)\phi_n}_{\text{reversal of incorrect decision in trial}} . \quad (1)$$

The left-hand side expression in the summation represents reversals that result from a type I error (false positive).[6] We can further simplify the relationship in equation (1) as follows:

$$\rho_n = q+(1-2q)\phi_n. \quad (2)$$

As seen, the relationship between reversals and welfare depends on whether q, which is unobservable, is greater or less than one half. In particular, it is impossible to conclude whether reversals reflect better or worse decision-making without further information about q. The contribution of this framework is that it allows us to infer the direction of the relationship without making any further assumptions.

To see this, recall that judicial panels are more likely than not to rule correctly, and are more likely to affirm than to reverse lower court decisions. This implies that cases reviewed by the high court are more likely to warrant affirmation than reversal. This result is formally summarized in the lemma below.

Lemma For any n, if $\phi_n > 1/2$ and $\rho_n < 1/2$, then $q > 1/2$.

With this result in hand, equation (2) implies that the coefficient on ϕ_n is negative and, as a result, higher reversal rates suggest more incongruence between the high court decisions and the law. This provides the link between the observable rate of reversals and a latent measure of social welfare generated by decision-making in high courts. This result is summarized in the following key proposition.

Proposition Given the assumptions, for any two judicial panels of size n and $n+1$,

$$\phi_n > \phi_{n+1} \Leftrightarrow \rho_{n+1} > \rho_n.$$

In particular, the optimal reversal rate ($\rho_n = 1-q$) is obtained when a judicial panel always rules correctly ($\phi_n = 1$), and approaches one half when decisions by the high court panel are random ($\phi_n = 1/2$). To see why more reversals indicate worse decision-making by the high court, notice what happens to the probability of a type I error in equation (1)

6. Figure 2 illustrates the possible outcomes on appeals. Papers using reversals as an outcome variable often assume that the reviewing court cannot make mistakes. For example, Alesina and Ferrara (2014) employ decision reversals to test a theory of racial bias in criminal sentencing. In contrast to my approach, they do not allow for false positives.

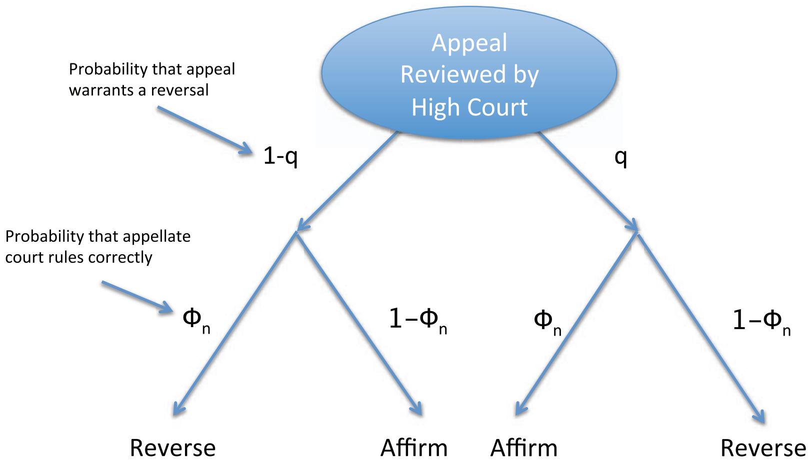

Figure 2. Paths to Outcomes on Appeals.
Notes: This chart shows the possible outcomes on appeals as a function of the probability that the law favors a reversal (q) and that the appeal court rules in a manner consistent with the law (ϕ_n). Appeals pending review are the set of appeals for which no further selection criteria are required and that are awaiting a formal review by the high court.

as ϕ_n falls. Because $q > 1/2$, higher rates of reversal more likely reflect false positives than corrections of lower court errors.[7]

5. Data

The data I use for my analysis come from multiple sources. Data on case-level state supreme court decisions are collected from the State Supreme Court Data Project. The dataset consists of all decisions made by state supreme courts between the years 1995 and 1998. There are over 20,000 court decisions on appeals over the course of these 4 years.

The unit of analysis is the decision of a state high court to affirm or reverse a lower court's decision upon reviewing an appeal.[8] Figure 3 illustrates the cross-sectional variation in panel size in the 50 states: 17 states have five high courts judges, 28 states have seven high courts judges, and 5 state supreme courts comprise nine judges each.[9] While there may be

7. The social benefit from reviewing the case stems primarily from the possibility of correcting an erroneous ruling by the lower court. The high court has more than one objective, such as reviewing appeals to clarify areas of the law and highlight a lower court's decision within a larger legal framework.

8. To be clear, the data include all appeals that were formally reviewed by the high court after being approved for review. In Section 8, I will describe the methods used for selecting appeals for review by state high courts.

9. Oklahoma and Texas have dual supreme courts, one for criminal appeal and one for civil appeal. In Texas, both have nine justices. In Oklahoma, designated as a panel of nine above, the supreme court has nine justices and the court of criminal appeals has five.

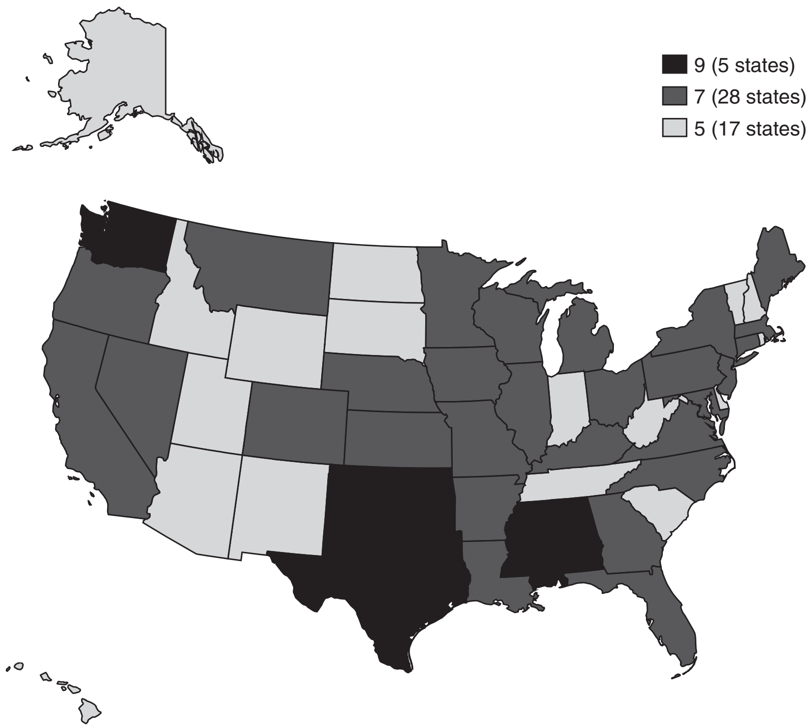

Figure 3. Map of State Supreme Court Panel Sizes.
Notes: This figure illustrates the present spatial variation in judicial panel size across state supreme courts. A darker color reflects a larger court. In Oklahoma, designated as a panel of nine above, the Supreme Court has nine justices and the (highest) court of criminal appeals has five.

incentives to change the size of high courts as a function of a state's political or economic environment, in practice, little if any such variation is observed. Panel sizes on each of the 50 US states' highest courts are generally fixed by state constitutions and are not easily modified.[10]

While the majority of appeals are reviewed *en banc*, there are a substantial number that are not, thus, providing within-state variation in the number of high court judges deciding a case. The reasons for this range from unexpected circumstances, such as death and mid-term departures

10. I found only four instances in which panel size was altered. In Virginia, a 1928 constitutional amendment resulted in an increase from five to seven state supreme court judges and the selection system was replaced from appointment to election. Arizona, Nevada, and Delaware passed amendments that increased the size of their supreme court from three to five justices in 1960, 1967, and 1978, respectively. And the most recent amendment to state supreme court panel size was Nevada's in 1997, which resulted in a second increase in panel size from five to seven judges.

Table 1. The Number of High Court Judges Deciding Appeal 1995–1998

	Panel5		Panel7		Panel9	
Panel composition	Decisions	Percent	Decisions	Percent	Decisions	Percent
en banc (entire court)	7392	82.47	9741	68.24	2655	89.76
en banc-1	1052	11.74	2225	15.59	62	2.10
en banc-2	519	5.79	2308	16.17	241	8.15
States	17		28		5	

Notes: This table tabulates the number and percent of state supreme court decisions between 1995 and 1998 by the number of judges present. Column headings indicate the number of judges who comprise the state high court; "*en banc*" is a decision made by the entire membership of the court. Decisions made not *en banc* can occur due to unexpected circumstances, such as deaths and mid-term departures from office. There are several cases in which more than two judges were missing; however, panels smaller than *en banc*-2 constitute a negligible number of instances and are omitted from analysis.

from office, to other institutional and administrative constraints. This in contrast to other court systems, such as the Canadian Supreme Court, in which panel size is endogenous.[11] Table 1 shows, for each panel size type, the number and percent of decisions made by the entire membership of the court as well as those made by incomplete panels. For example, 82% of decisions made in states that comprise five high court judges are made *en banc*, whereas 12(6)% of decisions are made in the absence of one (two) judge(s). Overall, in each of the three panel categories, at least 10% of decisions are not made *en banc*.

A second source for state supreme court decisions comes from Westlaw, an online resource for legal practitioners. The data include annual reports on the rate of reversals in state supreme courts between the years 2000 and 2011. Whereas the case-level dataset provides more detailed information on the cases and judges handling the case, it covers a shorter time period than the Westlaw dataset. Given that my objective is to capture the influence of court size on outcomes, a longer period of observation allows for greater judicial turnover; thus, exploiting the Westlaw data reduces the possibility that my findings are spurious or driven by a particular set of high court judges.

For robustness, I employ a set of case, panel, and state controls. I present summary statistics of variables I employ in the empirical analysis and their definition in Table 2. To control for judicial selection effects, I employ categorical variables for selection (four dummies) and retention methods (four dummies): legislative appointment, gubernatorial appointment, nonpartisan elections, or partisan elections. There are seven judicial panel demographics controls: means and standard deviations of age,

11. Since 2000, New Hampshire has used a fast-track to address appeals that are unlikely to generate legal precedent, where a sub-panel of three out of five high court judges are selected to review such cases. This change is not applicable to the case-level data that I use, which ends in 1998.

Table 2. Descriptive Statistics of Regression Covariates

Covariate	Panel5 Mean	Panel5 Standard deviation	Panel7 Mean	Panel7 Standard deviation	Panel9 Mean	Panel9 Standard deviation
Case characteristics						
Criminal	0.302	0.459	0.338	0.473	0.363	0.481
Civil government	0.289	0.454	0.323	0.468	0.215	0.411
Civil private	0.387	0.487	0.319	0.466	0.403	0.491
Juvenile	0.005	0.071	0.007	0.081	0.007	0.085
Trial court	0.887	0.317	0.816	0.387	0.916	0.278
Appellate court	0.003	0.056	0.044	0.206	0.002	0.044
Judicial selection						
Gubernatorial appointment	0.646	0.478	0.482	0.500	0.002	0.039
Legislature appointment	0.074	0.262	0.033	0.179	0.000	0.000
Nonpartisan elections	0.179	0.383	0.278	0.448	0.546	0.498
Partisan elections	0.059	0.236	0.205	0.403	0.450	0.498
Judicial retention						
Gubernatorial appointment	0.017	0.131	0.149	0.356	0.001	0.034
Legislature appointment	0.152	0.359	0.033	0.179	0.000	0.000
Nonpartisan elections	0.179	0.383	0.335	0.472	0.377	0.485
Partisan elections	0.073	0.260	0.123	0.328	0.450	0.498
Judicial panel covariates						
Proportion female (mean)	0.169	0.121	0.217	0.128	0.182	0.072
Proportion female (SD)	0.340	0.208	0.399	0.145	0.396	0.078
Age (mean)	57.494	4.437	58.170	3.917	55.216	5.269
Age (SD)	7.604	2.485	7.298	2.663	9.092	1.761
PAJID (mean)	39.721	17.570	43.300	11.856	33.523	9.317
PAJID (SD)	19.373	9.636	18.176	7.247	14.619	4.239
Herfindahl race index	0.900	0.151	0.831	0.165	0.833	0.081

Notes: These data come from the State Supreme Court Data Project covering decisions made by state high court judges serving between 1995 and 1998. Case characteristics and judicial selection and retention variables are indicator variables equal to one if observation is in specified category. For example, *Criminal* equals one if appeal is on criminal case, and zero otherwise. Panel characteristics variables were constructed using judge-level data; SD is standard deviation.

gender and PAJID scores (party-adjusted judge ideology), a common judicial ideological measure developed in Brace et al. (2000), and a Herfindahl index for race. These demographic covariates address the degree of within- and cross-panel heterogeneity as well as changes in panel composition. There are eight case dummy controls: general issue controls (criminal, civil government, civil private, juvenile, or non-adversarial) and court whose case is being reviewed (trial, intermediate, or state supreme court). For the state-level analysis, I impute panel characteristics and employ the same selection and retention variables; information on the cases is not available.

6. Empirical Strategy

To explore the relationship between high court judicial panel size and reversals of lower court decisions, I estimate regressions of the form:

$$\text{Reverse}_{in} = \alpha_0 + \alpha_1 \text{ Judges}_{in} + \mathbf{x}'_{in}\gamma + \varepsilon_{in}, \tag{3}$$

where Reverse_{in} is a dummy variable equal to one if a lower court's decision on case i was reversed by a panel of n judges, and to zero otherwise, Judges is a discrete variable that takes the number of high court judges who reviewed appeal i. Given the positive relationship between panel size and reversals, α_1 is expected to be positive. To examine the sensitivity of the estimates and to explore heterogenous selection effects, I include a vector $\mathbf{x}_{in}$ of case, state judicial selection (and retention), and panel demographics covariates described in the Data section. I also include year and state fixed effects and, unless noted otherwise, standard errors are adjusted for clustering at the state level.

Using data on decisions made *en banc* as well as the Westlaw data, I run regressions of the form:

$$\text{Reverse}_{is} = \beta_0 + \beta_1 \text{Panel7}_s + \beta_2 \text{Panel9}_s + \mathbf{x}'_{is}\delta + \eta_{is}, \tag{4}$$

where Panel7_s and Panel9_s are indicator variables for state s's high court panel size, and states with five high court judges are the omitted category. As illustrated in Figure 1, I expect the estimates for reversals to increase in panel size ($\beta_1 > 0$ and $\beta_2 > \beta_1$). The objective is to examine whether the differences highlighted in this figure are robust to controlling for additional covariates.

The main threat to identification is appeal selection. Specifically, if the appeals selected for review by larger panels are more prone to reversals, then the econometric model is misspecified. The best line of defense that addresses this concern is the use of within-court regressions, which include state fixed effects. Since appeals are typically reviewed *en banc*, any review by a sub-panel is plausibly unexpected and I do not consider appeal selection as a threat to identification in this set of regressions. The Appeal Selection section that follows further addresses threats to identification in the cross-court regressions.

7. Results

I first exploit the set of decisions made in sub-panels of high court judges. I report regression results in Table 3. The estimates for α_1 are statistically significant across specifications. The point estimate in the specification without controls (Columns (1)) is 0.0219 implying that adding one high court judge to the bench increases the likelihood of a lower court decision reversal by 2.2 percentage points. The estimate varies by less than 0.1 percentage points with the inclusion of year fixed effects, case and judicial selection covariates (Columns (2)–(4)). Adding the panel covariates reduces the estimates to 0.018 (Column (5)); however, in Column (6), with

Table 3. Reversals and the Number of Judges Deciding Appeals 1995–1998

Dependent variable: reversal of lower court decision						
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Number of judges[α_1]	0.0219***	0.0220***	0.0221***	0.0211***	0.0180***	0.0211**
	(0.00729)	(0.00731)	(0.00663)	(0.00656)	(0.00658)	(0.00817)
Year fixed effects		x	x	x	x	x
Case covariates			x	x	x	x
Selection covariates				x	x	x
Panel covariates					x	x
State fixed effects						x
Observations	26,848	26,848	26,848	26,848	26,714	26,714

Notes: This table presents OLS estimates for α_1 from equation (3). The unit of observation is appeal reviewed by state supreme court. Dependent variable is equal to one if lower court decision is reversed, and to zero otherwise. Number of judges is the number of state supreme court judges who reviewed the appeal. Case covariates are dummy variables for whether decision is on criminal, civil government, civil private, juvenile, or non-adversarial case, as well as indicators for court whose case is being reviewed: trial, intermediate, or state supreme court; selection covariates are indicators for selection and retention (each considered separately): legislative appointment, gubernatorial appointment, nonpartisan elections, or partisan elections. Panel covariates are means and standard deviations of judicial panel age, gender, and PAJID scores, a common judicial ideology measure, and a Herfindahl index for race. Standard errors are adjusted for clustering at the state level, * denotes 90% significance, ** denotes 95% significance, and *** denotes 99% significance.

the inclusion of state fixed effects, the estimated within-state marginal effect of one judge on reversals returns to 2.1 percentage points.

Next, I estimate equation (4) using only cases decided *en banc* in years 1995–1998. In the left panel of Table 4, I report OLS regression results.[12] For clarity, I present estimates for Panel7 and Panel9 in the top two rows. Below each specification, I provide three useful statistical entries:

1. the p value from a one-sided statistical test for a panel of seven resulting in more reversals than a panel of five judges ($\beta_1 < 0$),
2. the point estimate for the difference in reversals between a panel of nine and a panel of seven judges ($\beta_2 - \beta_1$), and
3. the p value from a one-sided statistical test for a panel of nine resulting in more reversals than a panel of seven judges ($\beta_2 < \beta_1$).

Since equation (4) allows for a non-monotone relationship between panel size and reversals, p values from one-sided tests can be used to infer whether or not such a relationship is plausible.

Overall, the signs of the estimates on panel size are positive as expected ($\beta_1 > 0$ and $\beta_2 > \beta_1$), with higher significance levels for Panel9 than for Panel7. The estimates in Table 4 suggest that, relative to panels of five judges (which reverse 28% of lower court decisions), panels of seven (nine) judges are 3–4 (13) percentage points more likely to reverse lower court decisions; the difference in reversals between panels of seven and panels of

12. For interpretation reasons, OLS regressions are easier to employ in a linear probability model such as the one I estimate, when the estimates do not fall beyond the unit interval. The results using probit regressions are consistent with those of Table 4.

Table 4. Reversals and Cross-State Variation in High Court Size

Dependent variable: reversal of lower court decision						
	Case-level data 1995–1998			State-level data 2000–2011		
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel7$^{[\beta_1]}$	0.0346	0.0365	0.0395**	0.0769**	0.0714**	0.0676**
	(0.0227)	(0.0230)	(0.0192)	(0.0314)	(0.0306)	(0.0322)
Panel9$^{[\beta_2]}$	0.133***	0.133***	0.125***	0.133**	0.116**	0.163**
	(0.0315)	(0.0373)	(0.0359)	(0.0521)	(0.0533)	(0.0684)
Year and region FEs		x	x		x	x
Controls			x			x
Observations	21,352	21,352	21,343	511	511	511
1. p Value, test $\beta_1 < 0$	0.0670	0.0590	0.0222	0.00891	0.0120	0.0206
2. Point estimate $\beta_2 - \beta_1$	0.0980	0.0965	0.0858	0.0564	0.0446	0.0953
3. p Value, test $\beta_2 < \beta_1$	0.00143	0.00539	0.0126	0.133	0.211	0.0521

Notes: This table presents OLS estimates for β_1 and β_2 from equation (4). The unit of observation is appeal reviewed *en banc* by state supreme court. Dependent variable is equal to one if lower court decision is reversed, and to zero otherwise. Panel7 (Panel9) is an indicator variable equal to one if the state high court comprises seven (nine) justices, and to zero otherwise; states with a panel of five high court justices is the omitted category. See list of controls in notes to Table 3. Regional dummies are Midwest, South, West, and Northeast. In state-level regressions, dependent variable is proportion of lower court decisions reversed by state supreme court per year, and panel controls are imputed from case-level data. Standard errors are adjusted for clustering at the state level, * denotes 90% significance, ** denotes 95% significance, and *** denotes 99% significance.

nine judges is 9–10 percentage points. In Column (1), the specification without controls, the point estimate for Panel7 is 0.0346 and for Panel9 is 0.133. The estimates obtained in Column (2) with the inclusion of year and region fixed effects are similar (0.0365 and 0.133 for Panel7 and Panel9, respectively). In the remaining specification, I include the full set of controls and fixed effects. The estimates are similar (0.0395 and 0.125 for Panel7 and Panel9, respectively) and statistically robust.

Using a third source of variation, I further corroborate my results using state-level annual data for years 2000–2011 and report regression results in the right panel of Table 4.[13] The estimates I obtain are consistent with those using case-level data. The coefficient estimates on Panel7 are larger than those I obtained in the case-level analysis, varying between 7% and 8% (versus 3%–4% using case-level data); it is possible that the estimates differ because the case-level dataset and state-level dataset cover different time periods (1995–1998 versus 2000–2011, respectively). The estimated marginal increase in reversals that results from a move to a panel of nine judges ($=\beta_2 - \beta_1$) is between 4% and 10% (versus 9–10% using case-level data). Overall, these results suggest that the effect of high court panel size cannot be easily explained away by the inclusion of covariates.

13. Since these data are given at the state-year level, case and panel covariates are not available. To impute panel controls, I merged state means from the case-level dataset.

Table 5. Reversals and Heterogeneity in Case Issue and Judicial Selection 1995–1998

Dependent variable: reversal of lower court decision						
	All cases		Criminal cases		Civil cases	
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Number of judges [α_1]	0.0206***	0.0226***	0.0168***	0.0383***	0.0214***	0.0211***
	(0.00315)	(0.00322)	(0.00568)	(0.00609)	(0.00374)	(0.00383)
Judicial selection	Appointed	Elected	Appointed	Elected	Appointed	Elected
Observations	14,394	12,320	4504	4165	9574	8013

Notes: This table presents OLS estimates for α_1 from equation (3). The unit of observation is appeal reviewed by state supreme court. Dependent variable is equal to one if lower court decision is reversed, and to zero otherwise. Number of judges is the number of state supreme court judges who reviewed the appeal. The full set of covariates is included in all specifications. See notes in Table 3 for list of covariates. Bootstrapped standard errors are adjusted for clustering at the state level, * denotes 90% significance, ** denotes 95% significance, and *** denotes 99% significance.

To explore heterogenous effects, I run regressions cutting the data by case issue and judicial selection. I report the results in Table 5. The results on panel size and reversals continue to hold in these subsamples of the data. There are two noteworthy observations. In states with elected judges, there are significantly larger effects of panel size in criminal cases than in civil cases; in states with appointed judges, the opposite is true but the difference is less noticeable. This implies the second observation: in states with elected judges the effects of panel size in criminal cases are larger than in states with appointed judges. To conclude, the effect of panel size on outcomes is evident in both criminal and civil cases, both under appointed and elected judges.

8. Appeal Selection

The results above suggest a positive relationship between reversals and panel size. This result cannot be explained away by the inclusion of a reasonably large set of covariates. I next collected additional data to examine three plausible channels, which I consider to be leading alternative hypotheses. Specifically, I assumed that, conditional on being accepted for review, the prior probability that an appeal is reversed is invariant to the number of judges on the reviewing panel (i.e., $q_n = q$). It is possible that at least cross-sectionally this is not the case. The key threat to the cross-sectional identification is that the prior increases with n. In this case, more reversals are expected in states with larger high courts simply because the proportion of accepted appeals that warrant reversals is higher, independent of panel size.

I present three empirical models of appeal selection, which offer an explanation for the relationship between panel size and reversals. These models differ in the particular mechanism each suggests but have in common the main conjecture that larger panels select appeals that are more likely to warrant a decision reversal.

The first mechanism is based on the presumption that reversing a lower court's decision is more resource intensive than affirming it. After all, it might require more legal justification to reverse a colleague than to affirm her decision. As a result, appellate courts might refrain from reversing a case even if they collectively conclude that a mistake was made in the trial. That said, as panel size increases the cost per judge decreases, making it less onerous for high court judges to reverse lower court decisions.[14] A second mechanism explores whether the appeal selection method employed is more stringent in states with larger high court judicial panels. If one of the high court's objectives is to correct erroneous decisions of lower courts, then a more rigorous procedure for selecting appeals will produce a higher proportion of dockets that warrant a decision reversal.[15] In particular, an appeal that results in affirmation can be viewed as an appeal that the high court erred in hearing. Thus, larger panels might be more effective in selecting appeals that more likely warrant reversal. A final, related, mechanism I explore is whether larger state supreme courts are of higher quality. If so, then larger panels are more likely to detect erroneous decisions in the appeal selection process and subsequently reverse them upon review.

8.1 Court Resources

To assess whether judicial resources are driving the results on case reversals, I collected data on court budgets, the salaries of Chief Justices and Associate Chief Justices, and number of clerks and research associates per judge. I was able to obtain precise and complete information on personnel for 37 states. In Table 6, I provide summary statistics of these data. Information on salaries come from the National Center on State Courts.

Personnel and court budget data come from multiple sources. In most cases, this information is available online, either on the state court's or the state government's website. The difficulty in obtaining these data on all states is because not all states specify the budget break-down within the court system (i.e., what fraction of the budget is allocated to lower courts versus the state supreme court). Budget figures were normalized to the 2012 fiscal year allocation to the state's supreme court. As for personnel, the median number of clerks per judge is two across high courts of varying panel size. Research Associates (also called Staff Attorneys in several states) run between one to two per judge.[16] There are, however, several outliers in the mix of clerks and other researchers. For example, California

14. A similar argument can be made if there is a social cost from a judge reversing the decisions of other judges. In this case, the larger the panel the less each judge can be held responsible for spearheading the reversal.

15. This relationship may also result from the previous mechanism I proposed, whereby more resources make it possible to employ more costly appeal selection methods.

16. I am less confident with these figures since they are somewhat subject to which employees our contact person perceived as research staff.

Table 6. State Supreme Court Size and Alternative Reversal Mechanisms

Court characteristic	Panel5 (1)	Panel7 (2)	Panel9 (3)
Court resources (median)			
Clerks (total)	10	14	18
Research associates (total)	4	6	13
Chief justice salary	$147,350	$156,957	$152,500
Associate chief justice salary	$145,350	$148,378	$150,000
Court budget	$6,004,100	$12,180,713	$9,972,843
Appeal selection (mean)			
Voteshare to grant appeal	0.35	0.46	0.48
No discretionary review	0.2	0.06	0
Mandatory petitions filed	399	464	607
Court quality (median)			
Citations (per judge-year)	13.24	14.035	10.83
Opinions (per judge-year)	23.31	23.385	20.64
Decisions followed	13	18.5	12
Overall court rank	18	26	43

Notes: Data on salaries come from the National Center on State Courts and were last updated in January 2012. Data on court budgets and personnel come from multiple sources, including phone conversations with state court officials and state court and government websites; budget values are reported in 2012 dollars. Data on resources of 13 state courts are incomplete: two of Panel5, nine of Panel7, and two of Panel9. See Table 8 for complete details on appeal selection methods. *Voteshare to grant appeal* is the fraction of votes required to grant appeal for review out of entire membership of court and is based on Table 8; *No discretionary review* is an indicator variable equal to one if appeal review by court is mandatory; *Mandatory petitions filed* is the average number of mandatory petitions filed per year in a given state high court. Data on citations and opinions come from Choi et al. (2008) and are based on years 1998–2000; data on the number of decisions followed by out-of-state courts are based on years 1986–2005 and come from Dear and Jessen (2007); *Decisions followed* indicates the number of high court decision followed by at least three out-of-state courts. Data on court rank come from the US Chamber of Commerce Institute for Legal Reform; these data are reported annually from 2002 to 2008 and biannually from 2009.

reported to have no clerks but a total of 91 Staff Attorneys. Similarly, Louisiana reported to have 40 researchers in total; thus, I cannot separate the number of clerks per judge from the number of other personnel.[17]

I present results using the balanced case-level dataset in Table 7. Because of missing data less than two-thirds of my sample has complete information on judicial resources. The full set of covariates and year and regional fixed effect are included in both specifications. In Column (1), I report estimates for Panel7 and Panel9 without any of the court resource covariates to form a benchmark for comparison, and the complete set of mechanism covariates are included in Column (2).

The estimates for panel size are significant across specifications. The estimate for Panel7 changes from 5.7 percentage points in Column (1) to 4.3 percentage points in Column (2), with the inclusion of the full set of resource covariates. Similarly, the estimate for Panel9 changes minimally

17. In such cases, where a court reported one figure for all research personnel, I labeled them as Research Associates.

Table 7. State Supreme Court Size and Alternative Reversal Mechanisms

	Dependent variable: reversal of lower court decision							
	Court resources			Appeal selection			Court quality	
Independent variable	(1)	(2)		(3)	(4)		(5)	(6)
Panel7[β_1]	0.0574**	0.0426*		0.0433**	0.0417*		0.0395**	0.0756***
	(0.0221)	(0.0247)		(0.0184)	(0.0218)		(0.0192)	(0.0234)
Panel9[β_2]	0.189***	0.192***		0.111**	0.111**		0.125***	0.179***
	(0.0429)	(0.0377)		(0.0455)	(0.0533)		(0.0359)	(0.0326)
Clerks		0.0469***	Acceptance threshold		−0.0778	Citations		0.00981***
		(0.0152)			(0.0893)			(0.00255)
Research associates		−0.000130	Mandatory review		−0.00980	Opinions		−0.00510***
		(0.00648)			(0.0681)			(0.000762)
Chief Justice salary		0.00127	Mandatory petitions		−3.73e−05	Decisions followed		−0.00514***
		(0.0228)			(3.13e−05)			(0.00170)
Associate Chief Justice salary		−0.00570				Overall court rank		0.000227
		(0.0261)						(0.000676)
Court budget		0.000781						
		(0.00101)						
Observations	12,722	12,722		17,321	17,321		21,343	21,343
1. p Value, test $\beta_1 < 0$	0.00753	0.0474		0.0119	0.0319		0.0222	0.00108
2. Point estimate $\beta_2 - \beta_1$	0.132	0.149		0.0678	0.0692		0.0858	0.103
3. p Value, test $\beta_2 < \beta_1$	0.00155	0.000102		0.0688	0.0762		0.0126	4.34e−06

Notes: This table presents OLS estimates for β_1 and β_2 from equation (4) using balanced data samples in each of the three panels: court resources, appeal selection, and court quality. *Clerks*, *Research Associates*, *Opinions*, and *Citations* are computed per judge. *Mandatory petitions* are computed per year. *Decisions followed* is the number of state supreme court decisions followed at least three times by out-of-state courts. *Chief Justice Salary* and *Associate Chief Justice salary* are given in $10,000s. *Court budget* is given in $100,000s per judge. See further details in Table 5. The full set of baseline covariates and year and regional fixed effects are included in all specifications; see notes in Table 3 for the list of covariates. Standard errors are adjusted for clustering at the state level, * denotes 90% significance, ** denotes 95% significance, and *** denotes 99% significance.

from 18.9 to 19.2 percentage points. At the same time, the number of clerks per judge has a positive and significant effect on reversals, as suspected. The point estimates indicate that the addition of one clerk per judge increases the likelihood of reversal by 4.7 percentage points. On the other hand, the salaries of the Chief Justice and Associate Chief Justice have insignificant and close to zero effects. Likewise, the court's budget does not influence the likelihood of reversals. Overall, panel size remains a significant correlate of reversals, which cannot be explained by variation in court resources.

8.2 Appeal Selection Method

The method used by state supreme courts to determine which petitions are granted for review varies across states. If methods that yield a greater proportion of dockets that warrant a reversal are (positively) correlated with panel size, then appeal selection method may shed light on my findings.[18]

In most cases, all members of the state supreme court vote on whether to review an appeal (i.e., the decision to review an appeal is made *en banc*). Typically, the variation across states in selecting appeals stems from the number of affirmative votes required. In Table 8, I tabulate the types of appeal selection method by state supreme court panel size. The modal rule employed is a majority of votes, with a majority of states requiring less than a majority to accept an appeal for review. In the empirical analysis that follows, I use three appeal selection covariates: the voteshare necessary to grant an appeal, a dummy for no discretionary appeal, and the number of mandatory petitions filed in a given year. The latter addresses a possible negative relationship between panel size and the number of petitions a court receives. For example, it may be more challenging to screen petitions for erroneous lower court decisions if the number of petitions is large. As a result, reversals may be more likely in high courts that receive a smaller number of petitions. Another potential appeal selection mechanism concerns the overall selectivity of the high court. That is, the fraction of discretionary appeals approved for review by the high court. Unfortunately, I was unable to obtain sufficient data on the total number of appeals to and approved by the high court to use in regression analysis. I present summary statistics of the covariates in Table 6.

I present regression results for the balanced panel in Table 7. The results are consistent with my previous findings and the estimates for appeal selection covariates are insignificant.[19] The coefficient estimates on Panel7 vary between 4.2 and 4.3 percentage points and those on Panel9 are

18. Since data on appeal selection methods were not publicly available, I contacted state court clerks and other public officials directly to solicit information on the process each high court employed to select the set of discretionary appeals it would ultimately review. Following multiple attempts, I was able to obtain this information in all but eight states.

19. Interestingly, the acceptance threshold requirement appears to be negatively correlated with panel size. That is, a more stringent acceptance rule is less likely to result in a reversal.

Table 8. State Supreme Court Appeal Selection Methods

Acceptance rule	Panel5 (1)	Panel7 (2)	Panel9 (3)
States with acceptance rule			
Majority of court	7	12	2
Majority−1	3	6	2
Majority−2[a]	1	2	—
One vote[b]	—	1	0
No discretionary review	2	2	—
Other[c]	1	1	—
N/A	3	4	1

Notes: This table presents the frequency of state supreme court appeal selection methods by panel size. "Majority" indicates that majority of high court members is required to support an appeal for further consideration.

[a]Three courts initially sit in majority subpanels. If broad consensus is not reached, then remainder of judges participate in *en banc* vote.

[b]Maine and New York grant criminal appeals with one vote.

[c]Idaho grants all appeals if filed within specified timeframe; otherwise, appeal is up for a vote. In Iowa, subpanels of judges reroute appeals to the intermediate or state supreme court for review.

identical 11 percentage points. In sum, I do not find evidence in support of appeal selection method as an underlying mechanism for my findings.

8.3 Court Quality

The final possibility for appeal selection that I evaluate empirically is whether court quality is correlated with court size; thus, more reversals may take place if appeal selection focuses more closely on identifying cases that warrant reversal. For this exercise, I obtained information from multiple sources. There is no consensus in the legal scholarship on what measure best reflects court quality; however, several measures, such as citations to court decisions, are used most prominently in the literature. In the regressions that follow, I employ data from Choi et al. (2008) on average citations and opinions per judge per year covering the period 1998–2000. To capture court quality as a whole, I use data on decisions followed by out-of-state courts. These data come from Dear and Jessen (2007) and include the number of state supreme court decisions that were followed at least three times by out-of-state courts between 1986 and 2005. I also employ a composite score for overall court rank provided by the US Chamber of Commerce Institute for Legal Reform. This score was reported annually from 2002 to 2008 and biannually from 2009.

In Table 7, I report regression results exploring the relationship between court size and quality. Overall, the estimates and statistical significance of panel size are comparable to the baseline results; thus, the quality measures do not undermine the influence of panel size on reversals. With the inclusion of all the quality controls, the estimates for panel size increase both in size and significance, and all quality estimates but the estimate for overall court rank are significant as well; opinions per judge and decisions followed by out-of-state courts are negatively correlated with reversals and citations per judge is positively correlated with reversals.[20] However, only opinions per judge is significant with and without the inclusion of other court quality covariates. I find that an increase in one opinion per high court judge per year is associated with a decrease of one-third to one half of a percentage point in case reversals.

In conclusion, the effect of panel size on reversals in state high courts is significant. This relationship is not explained by the appeal selection channels I explored.

9. Robustness

In Panel A of Table 9, I show how regression results vary across different outcome measures and subsamples of the data using the case-level and state-level datasets. Panel A reports regression results using the case-level data. In Columns (1) and (2), I present estimation results of equation (4) excluding cases that resulted in a partial reversal and partial affirmation and affirmation, respectively. In the former, the estimates are similar to those in the baseline results: about 4 percentage points more reversals occur in states with a high court panel of seven judges relative to five judges and another 13 percentage points on top of that occur in states with a panel of nine judges. The estimates for Panel9 are higher than in Table 4. Overall, excluding intermediate outcomes does not change the relationship between panel size and reversals. Comparing reversals to partial reversals in Column (2), I find that while strict reversals increase in panel size, the relationship is statistically weaker when omitting appeals that culminate in affirmation; specifically, the one-sided hypothesis for a panel of seven inducing more reversals than a panel of nine ($\beta_2 < \beta_1$) cannot be rejected. In Column (3), I show that partial reversals have no relationship with panel size. In particular, if appeals that culminate in partial reversals are those most complicated and challenging for the court to resolve, as suggested by Westergren (2003), then the results from Column (3) imply that high courts of varying sizes face similarly challenging appeals.[21] Finally, in Column (4), I find a negative relationship between affirmations and panel size, as expected.[22] In this specification, the importance of coding partial reversals as reversals is tested given that Reverse is an indicator equal to one only if the appeal results in a complete reversal. This also comports with the results in Column (3)

20. I also explore two other measures provided by Dear and Jessen (2007), the number of decisions followed at least once, and the number of decisions followed at least five times by out of state courts. The results are similar.

21. Alternatively, the effect of court size on the likelihood of a partial reversal is cancelled out by an opposite relationship between court size and the likelihood of granting an appeal that warrants a partial reversal.

22. Once partial reversals are excluded, the relationship between affirmations and panel size is precisely converse to that between reversals and panel size.

Table 9. Reversal Measure Robustness and Sample Selection

	Case-level data 1995–1998				State-level data 2000–2011			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Reverse	Reverse	Partial	Affirm	Reverse	Reverse	Partial	Affirm
Excluded outcomes	Partial	Affirm	None	None	Partial	Affirm	None	None
Panel7$^{[\beta_1]}$	0.0444	0.0422**	−0.0114	−0.0338	0.0762**	0.0412	0.0113	−0.0849**
	(0.0266)	(0.0198)	(0.0105)	(0.0257)	(0.0362)	(0.0550)	(0.0135)	(0.0319)
Panel9$^{[\beta_2]}$	0.178***	0.0696**	−0.000828	−0.154***	0.176**	0.0959	0.0118	−0.156**
	(0.0485)	(0.0301)	(0.0155)	(0.0454)	(0.0798)	(0.0980)	(0.0281)	(0.0696)
Observations	16,114	8732	18,149	18,149	511	511	503	511
1. p Value, test $\beta_1 < 0$	0.0506	0.0192	0.859	0.903	0.0201	0.229	0.202	0.995
2. Point estimate $\beta_2 - \beta_1$	0.134	0.0275	0.0106	−0.121	0.100	0.0547	0.000512	−0.0708
3. p Value, test $\beta_2 < \beta_1$	0.00571	0.132	0.232	0.992	0.0729	0.222	0.490	0.895

Notes: This table presents OLS estimates for β_1 and β_2 from equation (4). The unit of observation is appeal reviewed by state supreme court. Dependent variable is dummy (indicated above column number) for ruling on lower court decision; "Partial" ("Affirm") is indicator variable equal to one if lower court decision is partially reversed (fully affirmed). Panel7 (Panel9) is an indicator variable equal to one if the state high court comprises seven (nine) justices, and to zero otherwise; states with a panel of five high court justices is the omitted category. Excluded outcomes are subset of state supreme court decisions removed from sample. The full set of covariates and year and regional fixed effect are included in all specifications; see notes in Table 3 for list of covariates. Standard errors are adjusted for clustering at the state level, * denotes 90% significance, ** denotes 95% significance, and *** denotes 99% significance.

suggesting that treating partial reversals one way or another will not change the results since there appears to be no correlation between panel size and partial reversals.[23] In Panel B of Table 9, I report analogous regression results using the state-level dataset. Despite the relatively limited variation in these data, the estimates are less sensitive to coding decisions and sample selection as they apply to the relationship between panel size and reversals.

10. Conclusion

Much of the legal scholarship debates over and emphasizes how preferences may drive variation in judicial behavior and subsequent outcomes in the courts. I take an agnostic approach to preferences and focus instead on the implications of the hierarchical nature of court systems and the number of decision-makers employed at each level. I exploit variation in the number of judges on the 50 US state supreme courts and find that a marginal increase in the number of high court judges increases the likelihood that it will reverse a lower court's decision. This result does not appear to be driven by an omitted variable.

To offer an explanation for the empirical finding, I develop a simple framework that connects reversal and panel size with a natural measure of social welfare, the accuracy of decision-making in the high court. The theory suggests that based on the empirical findings increasing the number of decision-makers on a court may not be advisable. These conclusions are consistent with a broad set of findings suggesting adverse small group size effects on outcomes.

Conflict of interest statement. None declared.

Funding

I gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada.

References

Alarie, B., A. Green, and E. Iacobucci. 2011. "Is Bigger Always Better? On Optimal Panel Size, with Evidence from the Supreme Court of Canada," U Toronto, Legal Studies Research Paper No. 08–15.

Alesina, A. F., and E. L. Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing," 104 *The American Economic Review* 3397–433.

Alnuaimi, Omar A., Lionel P. Robert, and Likoebe M. Maruping. 2010. "Team Size, Dispersion, and Social Loafing in Technology-Supported Teams: A Perspective on the Theory of Moral Disengagement," 27 *Journal of Management Information Systems* 203–30.

23. I explored with an ordered choice model in which the dependent variable is ordinal and takes three values representing reversal, partial reversal, and partial affirmation and affirmation. The results do not shed further insight beyond that obtained by using the simpler binary specification for *Reverse*.

Austen-Smith, D., and J. S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem," 90 *American Political Science Review* 34–45.

Brace, P., L. Langer, and M. G. Hall. 2000. "Measuring the Preferences of State Supreme Court Judges," 62 *Journal of Politics* 387–413.

Cameron, C. M., J. A. Segal, and D. Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions," 94 *American Political Science Review* 101–16.

Choi, S. J., G. M. Gulati, and E. A. Posner. 2010. "Professionals or Politicians: The Uncertain Empirical Case for an Elected Rather than Appointed Judiciary," 26 *Journal of Law, Economics, and Organization* 290.

Choi, S. J., and M. Gulati. 2006. "Ranking Judges According to Citation Bias (As a Means to Reduce Bias)," 82 *Notre Dame Law Review* 1279.

Choi, S. J., M. Gulati, and E. A. Posner. 2008. "Judicial Evaluations and Information Forcing: Ranking State High Courts and Their Judges," 58 *Duke Law Journal* 1313.

Choi, Stephen J., Mitu Gulati, and Eric A. Posner. 2012. "What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals," 28 *Journal of Law, Economics, and Organization* 518–49.

Coles, Jeffrey L., Naveen D. Daniel, and Lalitha Naveen. 2008. "Boards: Does One Size Fit All?," 82 *Journal of Financial Economics* 329–56.

Condorcet, Marquis de (Marie-Jean-Antoine-Nicolas de Caritat). 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, l'Imprimerie Royale*. Paris: Imprimerie Royale.

Daughety, A. F., and J. F. Reinganum. 2000. "Appealing Judgments," 31 *Rand Journal of Economics* 502–25.

———. 2006. "Speaking Up: A Model of Judicial Dissent and Discretionary Review," 14 *Supreme Court Economic Review* 1–41.

Dear, J., and E. W. Jessen. 2007. "Followed Rates and Leading State Cases, 1940–2005," 41 *UC Davis Law Review* 683.

Drahozal, C. R. 1998. "Judicial Incentives and the Appeals Process," 51 *SMU Law Review* 469–1603.

Eisenberg, Theodore, Stefan Sundgren, and Martin T. Wells. 1998. "Larger Board Size and Decreasing Firm Value in Small Firms," 48 *Journal of Financial Economics* 35–54.

George, T. E., and C. Guthrie. 2008. "Remaking the United States Supreme Court in the Courts' of Appeals Image," 58 *Duke Law Journal* 1439.

Gerardi, D., and L. Yariv. 2008. "Information Acquisition in Committees," 62 *Games and Economic Behavior* 436–59.

Gershkov, A., and B. Szentes. 2009. "Optimal Voting Schemes with Costly Information Acquisition," 144 *Journal of Economic Theory* 36–68.

Hall, M. G. 1992. "Electoral Politics and Strategic Voting in State Supreme Courts," 54 *Journal of Politics* 427–46.

Hall, M. G., and C. W. Bonneau. 2006. "Does Quality Matter? Challengers in State Supreme Court Elections," 50 *American Journal of Political Science* 20–33.

Huber, G. A., and S. C. Gordon. 2004. "Accountability and Coercion: Is Justice Blind When It Runs for Office?," 48 *American Journal of Political Science* 247–63.

———. 2007. "Directing Retribution: On the Political Control of Lower Court Judges," 23 *Journal of Law, Economics, and Organization* 386.

Huther, Jeff. 1997. "An Empirical Test of the Effect of Board Size on Firm Efficiency," 54 *Economics Letters* 259–64.

Iaryczower, M., G. Lewis, and M. Shum. 2013. "To Elect or to Appoint? Bias, Information, and Responsiveness of Bureaucrats and Politicians," 97 *Journal of Public Economics* 230–4. Amsterdam: Elsevier.

Iaryczower, M., and M. Shum. 2012. "The Value of Information in the Court: Get It Right, Keep It Tight," 102 *American Economic Review* 202–37.

Koriyama, Y., and B. Szentes. 2009. "A Resurrection of the Condorcet Jury Theorem," 4 *Theoretical Economics* 227–52.

Kornhauser, L. A., and L. G. Sager. 1986. "Unpacking the Court," 96 *Yale Law Journal* 82–117.

Lim, C. S. H. 2013. "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges," 103 *American Economic Review* 1360–97.

Mak, Yuen Teen, and Yuanto Kusnadi. 2005. "Size Really Matters: Further Evidence on the Negative Relationship Between Board Size and Firm Value," 13 *Pacific-Basin Finance Journal* 301–18.

Martinelli, C. 2006. "Would Rational Voters Acquire Costly Information?," 129 *Journal of Economic Theory* 225–51.

Mukhopadhaya, K. 2003. "Jury Size and the Free Rider Problem," 19 *Journal of Law, Economics, and Organization* 24.

Palfrey, T., and H. Rosenthal. 1984. "Participation and the Provision of Discrete Public Goods: A Strategic Analysis," 24 *Journal of Public Economics* 171–93.

Persico, N. 2004. "Committee Design with Endogenous Information," 7 *Review of Economic Studies* 165–91.

Petty, Richard E., Stephen G. Harkins, Kipling D. Williams, and Bibb Latané. 1977. "The Effects of Group Size on Cognitive Effort and Evaluation," 3 *Personality and Social Psychology Bulletin* 579–82.

Posner, R. A. 1998. *Economic Analysis of Law*, Vol. 5 New York: Aspen Law & Business.

———. 2005. "Judicial Behavior and Performance: An Economic Approach," 32 *Florida State University Law Review* 1259–415.

Shavell, S. 1995. "The Appeals Process as a Means of Error Correction," 24 *Journal of Legal Studies* 379–426.

Spitzer, M., and E. Talley. 2013. "Left, Right, and Center: Strategic Information Acquisition and Diversity in Judicial Panels," 29 *Journal of Law, Economics, and Organization* 638–80.

Watson, A. S. 1987. "Some Psychological Aspects of the Trial Judge's Decision-Making," 39 *Mercer Law Review* 937.

Westergren, S. 2003. "Gender Effects in the Courts of Appeals Revisited: The Data Since 1994," 92 *Georgetown Law Journal* 689.

Wheelan, Susan A. 2009. "Group Size, Group Development, and Group Productivity," 40 *Small Group Research* 247–62.

Yermack, David. 1996. "Higher Market Valuation of Companies with a Small Board of Directors," 40 *Journal of Financial Economics* 185–211.