Optimal panel size on the UK Supreme Court

Chris Hanretty
University of East Anglia

February 17, 2016

Abstract

I investigate the determinants of panel size on the UK Supreme Court. The size of the panel in each case is decided by the most senior judges on the court, aided by court staff. Panels may be formed of five, seven, or nine judges. The court says that larger panels will be formed in cases of great constitutional or public importance. I test this claim alongside (i) a theory of optimal panel size derived from Condorcet's jury theorem; and (ii) a theory of panel size based on the need to minimize the risk of government non-compliance. There is some support for the theory of optimal panel size, but no support for the theory that courts choose larger panels if they fear non-compliance.

Note: This paper was written as part of a book-length project on decision-making in the UK Supreme Court. I apologise in advance for any consequent lack of context.

Optimal panel size

Introduction

In 1785, the Marquis of Condorcet published his eponymous jury theorem (Condorcet 1785, Condorcet and Baker (1976)). The theorem concerns a group which must decide, by majority vote, between two alternatives: one right, one wrong. Condorcet showed that if members of the group are just fractionally more likely to pick the "correct" alternative, then the larger the group, the more likely they are to decide correctly.

Although in other respects inaptly named, the theorem describes well the jury's task. Most legal systems only allow juries two options: to convict or to acquit. (The "not proven" verdict possible in Scots criminal law is a notorious exception). In most criminal cases, it is clear that *either* the defendant did do the things of which they were accused, or they did not do those things. If we gave as much weight to not freeing the guilty as we do to not convicting the innocent, Condorcet's theorem would provide a solid basis both for majority decision-making on juries, and for larger juries wherever possible.[1]

The jury theorem can be modified to account for more than two options (List and Goodin 2001), but it does seem to require that one option be right, or better, and others wrong, or inferior. It would be difficult to see any point

[1] In practice, we care more about convicting the innocent than we do about letting the guilty go free. As a result, juries are often required to decide by unanimity or super-majority.

in the theorem otherwise. There is a lively philosophical debate about whether there are always right answers to hard legal questions (Dworkin 1985, chap. 5). My aim in this paper is not to prove that there are always right answers, and that therefore Condorcet's jury theorem must also apply to panels of judges. Rather, I aim to use Condorcet's jury theorem to elucidate the Supreme Court's decisions regarding the size of the panel which must hear each case.

The Supreme Court's rules regarding panel size are set out in statute and online. Article 42 of the Constitutional Reform Act of 2005 states that the Court is "duly constituted in any proceedings" if and only if it consists of an uneven number of judges greater than or equal to three.[2] Neither the Rules of the Supreme Court nor its Practice Directions provide any further detail on panel size, but the court website does list criteria "to be used when considering whether more than five Justices should sit on a panel":[3]

- If the Court is being asked to depart, or may decide to depart from a previous decision.
- A case of high constitutional importance.
- A case of great public importance.
- A case where a conflict between decisions in the House of Lords, Judicial Committee of the Privy Council and/or the Supreme Court has to be reconciled.
- A case raising an important point in relation to the European Convention

[2]Here, "proceedings" must be interpreted to mean any case heard and decided by the Supreme Court. Otherwise, the court would not have been duly constituted in the permission to appeal hearings for Anoliefo (Appellant) v Her Majesty's Advocate (Respondent) (Scotland) (Case no. UKSC 2014/0131), Kapri (Appellant) v Her Majesty's Advocate (Respondent) (Scotland) (Case no. UKSC 2014/0131), Sabiu (Appellant) v Procurator Fiscal, Fort William (Respondent) (Scotland) (Case no. UKSC 2014/0134), and O'Leary (Appellant) v Her Majesty's Advocate (Respondent) (Scotland) (Case no. UKSC 2015/0150), which were all heard by panels of four judges.

[3]These criteria can be found at "Panel numbers criteria", https://www.supremecourt.uk/procedures/panel-numbers-criteria.html, last accessed 2016–02–01.

on Human Rights.

Though the Constitutional Reform Act permits panels of three and eleven judges, the former has been used sparingly and the latter not at all. Three judge panels handed down judgments in *Anderson v Shetland Islands Council and another* ([2012] UKSC 7) and *Apollo Engineering Limited v James Scott Limited* ([2013] UKSC 37). Both cases concerned questions of procedure brought by "determined" litigants in person: security for costs in *Anderson*, and the possibility of appeal to the Supreme Court in *Apollo Engineering*. Though an eleven judge panel is theoretically possible, "asking 11 to sit would have left the remaining Law Lord feeling somewhat excluded" (Dickson 2007, 592). If we are prepared to ignore these rarer possibilities, then the question then becomes: what factors determine whether the Supreme Court sits in a panel of five judges, as opposed to a panel of seven judges or nine judges? One factor, distinct from the criteria set out above, relates to the difficulty of deciding the case, and employs Condorcet's jury theorem to surprising effect.

Some theory

Getting it right at the margins

The intuition behind Condorcet's jury theorem is simple. Voters have a chance (or probability) of getting the right answer. This probability can be higher or lower. If this probability is greater than 0.5, a single voter is more likely than not to arrive at the right answer. However, since individual voters are fallible, a single voter may reach the wrong answer in any given case.

When we move from one to three voters, all three voters remain just as they were before: more likely than not to reach the answer, yet still fallible. However, it would be unfortunate in the extreme if all three were to fail to reach the right

answer. As we increase the number of voters, then eventually the *proportion* of voters reaching the right answer will equal the *probability* that each individual voter reaches the right answer. Where that probability is greater than 0.5, this means that as we add more voters, the proportion of voters will also end up being greater than 0.5, and a proportion of voters greater than 0.5 is a majority.

Though the intuition is simple enough, it is far less straightforward to calculate the probability that a group of n voters deciding by majority vote will reach the right result.[4] Accordingly, Figure 1 shows the probability that groups of 3, 5, 7, 9, and 11 members will reach the right decision when each individual member of the group decides with probability p. I have chosen three different probabilities (0.55, 0.66, 0.9), which we might interpret either as the probabilities of average, good and expert judges, or as probabilities in difficult, moderate, and easy cases. For reasons that will become clear, I will adopt this second interpretation.

What is not clear from the upper panel of Figure 1 is the additional benefit that comes with increasing group size. This is plotted in the panel below. For all three probabilities, the change in the *group* probability of reaching a correct decision which results from adding two more panel members — the *marginal benefit* — decreases as the panel becomes larger. However, the marginal benefit is greatest where cases of intermediate difficulty ($p = 0.66$). The marginal

[4] If each voter gets the right answer with probability p, then the group probability is equal to:

$$P_n = \sum_{m=\frac{n+1}{2}}^{n} \frac{n!}{m!(n-m)!}\, p^m (1-p)^{n-m}$$

This formula is split into three parts. Reading left to right, these are summation; enumeration of possibilities; and multiplication of probabilities. The summation operator says that the steps that follow will be calculated for progressively larger majorities of right-voting group members, starting from the smallest possible majority ($m = \frac{n+1}{2}$) and ending with all members of the group voting the right way. The second stage, the enumeration of possibilities, calculates the number of combinations of n group members taken m at a time. The final stage multiplies the probability of a correct decision by itself as many times as there aare members of the putative majority, and then multiplies this by the complementary probability, multiplied by itself as many times as there are members of the putative minority.

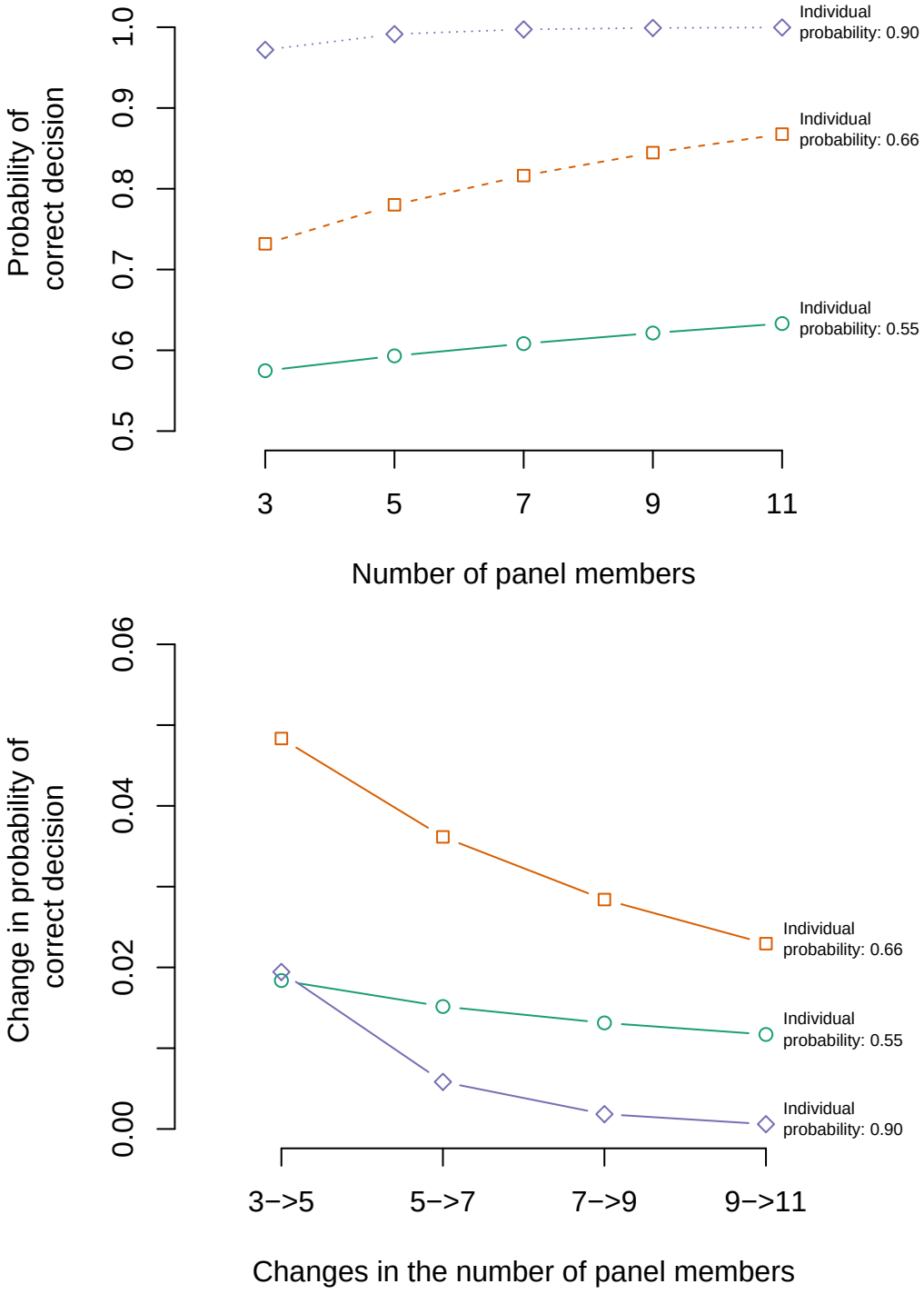

Figure 1: Committee size and correct decisions

benefit is generally lowest in the easiest cases, but the marginal benefit can also be low for very difficult cases.

This makes sense. In very easy cases, we have some confidence that a small panel of three or five judges will reach a right decision: anything else would be overkill. Conversely, in very difficult cases, where judges are only slightly better than a coin-toss ($p = 0.55$), adding more judges is going to deliver a slight benefit, but it would take a cast of hundreds to make that slight difference in accuracy tell.

The marginal benefit of two additional judges thus changes according to the difficulty of the case. If the Supreme Court had unlimited resources, these changes in marginal benefit would not matter. Since additional judges always make a correct decision more likely, panels on an unconstrained court should always be as large as possible.

In practice, appointing further judges to a panel comes at a cost. This cost is essentially an opportunity cost. Two judges who are added to a panel will have less time to sit on and decide other cases. While it would be possible for the Supreme Court only ever to sit in panels of nine (thereby maximizing the chances of a correct decision according to the jury theorem), this would mean that the Supreme Court would be forced either to hear fewer cases, or spend less time deciding each case, or build up a considerable and ever-increasing backlog of cases to be decided. As one former Lord Justice of Appeal put it, "unless the Supreme Court takes a radically different view... as to the number of appeals that it needs to hear and the amount of its time that it permits counsel to occupy in adding to their extensive written cases by oral argument, to sit in larger panels with any sort of regularity would cause the list to slow up to an entirely unacceptable degree" (Buxton 2009, 290)

If we are prepared to accept that this opportunity cost is constant — that

adding two judges to a five-judge panel "costs" the court as much as adding two judges to a seven-judge panel — then we reach a counter-intuitive conclusion: the largest panels should be reserved for cases of intermediate difficulty, since it is in these cases that the marginal benefit is the greatest, and thus most likely to outweigh the opportunity cost of adding judges. If the President were to assign a nine-judge panel to an extremely difficult case, he would be empanelling two judges who would contribute relatively little to the panel's probability of reaching a correct decision, and whose time and energy could be better used elsewhere.

In suggesting this, I am following the reasoning laid out in Alarie, Green, and Iacobucci (2011), who explore panel size on the Supreme Court of Canada which, like the UK Supreme Court, sits in panels of five, seven, or nine judges. Alarie, Green, and Iacobucci (2011) attempt to formalize the net benefit of five-, seven- and nine-judge panels. Simplifying slightly,[5] the net benefit of sitting in a panel of five judges is equal to the probability that a five-judge panel will reach the right decision, multiplied by the importance of the case, minus the opportunity cost of the five judges so empanelled.

Alarie, Green, and Iacobucci (2011) have no measure of the probability that panels of different sizes will reach different decisions, and so are forced to proceed by examining things from both ends of the telescope: first assessing whether the panel was "efficiently-sized" by examining voting behaviour; and then examining voting behaviour by looking at panel size. Below I set out a measure of the ex ante probability of a correct decision, which builds on a measure of "opinion below". Before I do that, I set out two alternative accounts of the choice of panel size: one which is based on what the court and its judges say, and one which is based on the risk that the court may have to rule against a government

[5] Alarie et al. add an additional parameter, which reflects the probability that the lower court decision will be correct.

which can threaten to evade the court's decision.

What the Court says, and what the judges say

In the preceding section, I set out an interpretation of the Supreme Court's decision to hear cases in panels of different sizes. That interpretation used a formal technical language which would not, I suggest, be used by the Court or by individual judges. Although there is no obligation to explain organizational behaviour in terms familiar to its participants, there is a presumption that individuals' own accounts of what they thought there were doing have some probative value. It is therefore helpful to set out what the Court says it is doing when it decides to hear cases in differently-sized panels, and what individual judges say they are doing — for, as I shall show, these two sets of accounts point in different directions.

Any references to the *Court* deciding panel sizes are a convenient shorthand for a more complex decision-making process. Panels are drawn up by the Registrar of the Court, who performs this task for both the Court and the Judicial Committee of the Privy Council. The Chief Executive of the court helps draft these lists (Darbyshire 2011, 370).[6] These lists are then put to the President and Deputy President of the court, who have "ultimate authority" and who may make alterations (Paterson 2013, 72).

The President (and to a lesser extent, the Registrar: Paterson (2013), 72–73) may be lobbied by colleagues who want to sit on a particular case, and successful lobbying may lead to the panel being enlarged. Darbyshire (2011) suggests that *Radmacher v. Granatino* ([2010] UKSC 42) was originally listed to be heard by a panel of seven judges, but was increased to nine judges as a result of successful lobbying from one of the judges not chosen. The successful lobbyist reported

[6] Louise di Mambro has been Registrar since the Court was established. The Chief Executive for most of this period was Jenny Rowe, who was replaced in September 2015 by Mark Ormerod.

that "the President is infinitely malleable on size and constitution of panels" (Darbyshire 2011, 372).[7]

Infinitely malleable though he may be, the President is in theory guided by the criteria set out at the beginning of this paper, which state that the court will hear cases in larger panels in cases of high constitutional or great public importance, and/or which involve human rights complaints, a challenge to established precedent, or conflict between courts. These criteria, which have been criticised for their vagueness (Knight 2012, 477), cannot have very much bite: they postdate the court's first forty decisions, and were only written and placed on the court's website after participants at a seminar to mark the court's first anniversary suggested such guidance might be useful.[8]

The suggestion that the Court might be solely guided in its decisions by its published guidelines is also undercut by judges' own remarks on the subject. These remarks concern not just judges' general preferences for larger or smaller panels (they are generally in favour of larger panels: Darbyshire (2011), 372), but the reasons they give for sitting in larger panels in particular cases. Judges seem very conscious that particular cases are more likely than others to result in 'bare majority' decisions (Darbyshire 2011, 373; Paterson 2013, 72*fn*35). 'Bare majorities' are particularly troubling in cases heard by five judges, since these cases raise the possibility that the outcome of the case would have been different had the case been heard by a different group of judges. Although 'bare majorities' are no less likely in panels of seven or nine judges — indeed, they may be more likely — they are less embarassing: it is less likely that a different group of nine judges would decide differently, simply because a different group of nine judges would perforce include many of the same judges who originally

[7]At the time, Lord Phillips was the President of the Court. Other Presidents may prove to be less malleable.

[8]The issue subsequently came up in meetings of the Supreme Court/Judicial Committee of the Privy Council "User Group": see the minutes of the meeting of the 25th January 2013

decided the case.

This concern is not inconsistent with the Court's published guidelines. Indeed, it would be perfectly consistent if cases likely to result in bare majorities were also cases likely to raise human rights issues, or which were otherwise of high constitutional or great public importance. At the same time, however, this concern also speaks to some of the issues raised above in the discussion of the jury theorem. A case likely to result in a bare majority is also a case in which the *ex ante* probability of judges arriving at the correct decision is low. In such cases, judges' remarks and the jury theorem point in opposite directions: the jury theorem suggests that the panel should be small (because the marginal benefit of another judge is outweighed by the opportunity cost), whilst judges remarks suggest that the panel should be larger (because a larger panel, though it may or may not be better placed to identify the right decision, is less likely to lead to embarassment). We are in the rare position of having two commensurable theories of panel size which make sharply opposing predictions.

Larger panels as a means of drawing attention to cases

Courts sometimes require governments to do things they do not want to do, or to refrain from doing things they would rather continue doing. At such times, governments face a choice between complying with the court's decision and evading it.

This is a genuine choice, even if the second option of evasion is rarely chosen. In the UK, the government has in practice chosen to evade the European Court of Human Rights' decision in *Hirst v United Kingdom (No 2)* [(2005) 42 EHRR 849]. The UK government's continuing non-compliance with *Hirst* has caused the ECtHR, on one view, to adopt more "politically sensitive" positions in subsequent prisoner voting rights cases (Briant 2011, 246). Even where evasion

does not take place, the threat of evasion can cause courts to limit their decisions to within a politically acceptable range (Stone 1992).

The fact that we do not see widespread non-compliance therefore stands in need of some explanation.[9] One particular group of explanations focuses on the role played by transparency and public awareness of the court (Vanberg 2004; Staton 2006). If governments risk a public backlash for evading court decisions, they may choose compliance; but public backlash can only arise if the public is aware of a decision. Since most courts of last resort enjoy greater public support than most national governments, courts have a standing incentive to maximize the transparency of their decisions in order to forestall possible non-compliance. Measures to increase transparency may have a cost, and so may not always be employed, but may instead be reserved for cases in which the risk of non-compliance is greatest.

This argument from transparency can sound as though it imposes great demands on the public. Certainly, if widespread public awareness of specific cases decided by courts of last resort were a necessary condition for compliance, evasion would be the norm. But backlash may be confined to a very few people whilst still acting as a "powerful enforcement mechanism... politicians and party leaders are concerned about shifts in support *at the margin*. Thus, to be effective in deterring evasion, the mechanism does not require that all, or even many, citizens are motivated by concerns for judicial authority" (Vanberg 2004, 20).

Instead of seeking to explain compliance by way of transparency, we can run the causal chain in reverse, and try and explain transparency-promoting measures by pointing to the risk of non-compliance. This has led to investigations of how courts selectively promote some of their case results (Staton 2006), or how

[9]The fact that non-compliance has been "treated as a peripheral, implementation problem" (Cooper 1996, 255) by British public lawyers may be an oblique tribute to the British state.

they hold oral hearings when oral hearings are not the norm (Vanberg 2004, chap. 4).

This leads to a consideration of panel size. Whether or not larger panels have desirable epistemic properties, they are an important signal of case importance. Cases heard in larger panels generally attract more media attention than cases heard in panels of five judges. In making its decisions about panel size, the court is therefore making a decision which has consequences for the level of public awareness surrounding each case.

If the court therefore wishes to attract extra attention to the case, perhaps because it fears that the government will drag its feet in complying, it should therefore sit in a larger panel. In practice, this reasoning suggests that the Court should sit in larger panels in

- cases involving central government;
- cases where the court believes its judgment may go against the government;
- cases involving human rights claims

Human rights claims matter for panel size because of the particular way in which human rights claims are dealt with in the UK. Where legislation is not compatible with individuals' human rights, and cannot be interpreted in a way that would make it compatible, courts may issue a declaration of incompatibility. Legislation which has been declared incompatible remains in force until it is altered by the government. At the time of writing, the only declaration of incompatibility which has not been acted upon is the declaration of incompatibility made by the Scottish Registration Appeal Court in *Smith v Scott* (2007 SC 345), which followed on from the ECtHR's decision in *Hirst* with which I began this section. The requirement that the government remedy situations of incompatibility thus poses an additional risk of non-compliance.

Testing these accounts

Above I set out three accounts of the choice of panel size: one account based on the jury theorem; one account based on what courts and judges say; and one account based on the risk of non-compliance. These accounts identified different factors which should affect the choice of panel size. Here, I describe these factors and their operationalization in more detail.

Measurement

The first factor identified was the **probability of reaching the correct decision**, which featured in the account based on the jury theorem. This is an *ex ante* probability, both in the sense that it cannot be inferred from the actual decision the court took, and in the sense that it precedes other steps which the court might have taken to improve its chances of deciding correctly (by, e.g., asking questions of counsel, reading and re-reading the case materials, deliberating, and so on). In theory the probability can vary across judges as well as cases. Some judges might be more likely to be right than other judges, or more likely to be right in particular cases. I ignore this variation, and concentrate on how this probability varies across cases.

I base my measure of the probability of reaching the correct decision on the measure of opinion in the courts below. Most cases do not arrive at the Supreme Court as *tabulae rasae*. Between one and six judges will have heard the case, and found either for the appellant or the respondent. In some cases, all of these judges will have found in favour of the respondent. In other cases, the proportions will be more balanced. The most favourable proportions for appellants occur when the appellants won at first instance before losing on appeal before a court in which there was a dissenting opinion. Since cases at first instance are typically heard by a single judge, and cases on appeal are typically heard by a

panel of three judges, the percentage of judges who had decided in favour of the appellants is 50%. This is the most favourable proportion for appellants for the simple reason that if the appellants had convinced more judges, they would no longer be the appellants but would instead be the respondents.

This simple proportion is an intuitive measure of "opinion below". But a simple proportion does not capture all aspects of opinion below. Most crucially, it ignores the weight of opinion on either side.

Consider two hypothetical appeals before the Supreme Court. In one appeal, the appellants appeal from a decision of the Court of Appeal, having first lost in the High Court. Thus, all four judges who have heard the case have found against the appellants, and the proportion of judges who have heard in their favour is zero percent. In a second appeal, the appellants appeal from a decision of the High Court. The sole judge who has heard the case has found against the appellants, and thus the proportion of judges who have found in favour of the appellants is also zero. Although the percentages are the same in both cases, we would ordinarily think that the appellants in the second case were more likely to succeed, needing only to overturn the opinion of one judge.

One way of dealing with scenarios like these is to use Laplace's Rule of Succession, a rule for estimating the subjective probability to assign to one of two possible events given a finite number (n) of previous observations and a number of "successes" ($s, s < n$). The rule emerges naturally from a prior belief that either event is equally possible. As Wikipedia puts it, this prior belief is "as if we had observed one success and one failure for sure before we even started the experiments. In a sense we made $n + 2$ observations (known as pseudocounts) with $s+1$ successes". The (subjective) probability we should assign to the event is thus:

$$p = \frac{s+1}{n+2} = \frac{0+1}{0+2} = \frac{1}{2} = 50\%$$

If we observe one more "success", the probability we assign rises, but is still some way short of the observed proportion of successes: instead of revising our probability to 100% (completely certain), it moves only to $\frac{2}{3} = 66.67\%$. As we observe a constant stream of successes, our subjective probability approaches, but never reaches, 100%.

The rule of succession helps us deal with the two hypothetical cases described above. It does so not by providing us with numbers to assign as subjective probabilities — the rule of succession assumes away all other sources of prior information, and the point of this book is to marshal such information — but rather by providing a single number which encapsulates the weight of opinion below. In the first case — the case which was heard by four judges — we can form a subjective probability of the appeal succeeding by dividing one plus the number of judges who found in favour of the appellants (zero) by two plus the number of judge who heard the case (four). Our subjective probability thus becomes $\frac{1}{6}$, or 16.67%. In the second case, heard by only one judge, the probability of success is much higher, at one in three.

Heard the case	**Found for the appellants**	**Measure**	**# cases**
6	0	0.12	5
4	0	0.17	267
3	0	0.20	14
2	0	0.25	18
6	1	0.25	1
1	0	0.33	10

Heard the case	**Found for the appellants**	**Measure**	**# cases**
4	1	0.33	160
3	1	0.40	3
2	1	0.50	6
4	2	0.50	22

The table above shows the values of this measure for all cases decided by the Supreme Court, excluding those cases which arrived at the court as reference questions. The same value of the measure can correspond to one or more appellate histories. Thus, one judge may find for the appellant, but this may be a judge in the Court of Appeal who dissents, or a judge of the High Court who is over-ruled. Additionally, because of the way the rule of succession works, a case can be assigned a value of 0.33 when a single judge finds against the appellants, or when three of four judges do so.

This measure of opinion below runs from a theoretical lower bound of 0% to a maximum of 50%; the former value represents an entirely unanimous judicial hierarchy; the latter a very divided one.

If we are prepared to accept that judges at all levels of the hierarchy are more likely than not to decide cases correctly, and that judges in the Supreme Court are comparable to judges at other levels of the judicial hierarchy, then this measure of opinion below should be inversely associated with the ex ante probability of a correct decision. The greater the value of this measure, the more divided the judicial hierarchy, and thus the lower the prospect that any individual judge of the Supreme Court will decide the case correctly. After all, if two judges found for the appellant and two judges found for the respondent, then the *observed* probability that some judges have got it wrong is fifty percent.

Since our measure of the probability must, by assumption, run between slightly more than fifty and one hundred percent, then we must adjust the measure of opinion below. We start at one hundred percent, and subtract our measure of opinion below, and add on one to ensure our measure is strictly greater than fifty percent:

$$Prob(\text{correct decision}) = 100 - \text{Opinion Below} + 1$$

which in practice gives us probabilities which range from 0.51 to 0.885.

The second factor alluded to in the discussion of the jury theorem is the cost of judges' time. In Alarie, Green, and Iacobucci (2011), the opportunity cost of two additional judges' time disappears from the empirical analysis, because it is assumed to be constant; and since it is constant in the decision to move from five to seven judges in the same way as it is in the decision to move from seven to nine judges, it can be simplified away. The same is true in the statistical model I use (and which I describe below): there is no distinct parameter which corresponds to the cost of two more judges' time.

One variable which I do use, however, does track changes in that opportunity cost — the marginal cost at the margin, so to speak. That variable is the court's **workload**. The expectation is that the greater the court's workload, the more assigning additional judges to hear the case distracts judges from finishing other cases.

I suggest that the workload of individual judges can be measured by counting **the number of cases which they have heard, but which have not yet been decided**, including cases heard by the Judicial Committee of the Privy Council alongside cases heard by the Supreme Court. In effect, I equate workload with the size of the judge's in-tray. In proposing this as a measure, I am ignoring all other sources of work, including but not limited to participation

in the permission to appeal process, administrative tasks on the court, and all forms of public engagement. That does not mean these things are unimportant, just that they are secondary. Although more senior members of the court are disproportionately likely to carry out these other activities, they also tend to have higher caseloads.

This is a simple measure, and perforce crude. Some cases are simply more onerous than others. Writing an opinion in a case is more onerous than concurring in a colleague's opinion. Yet given that I aim to explain outcomes such as opinion writing, the benefits of a simple measure seem to outweigh the costs.

Here I explain panel size with reference to the workload of the court, rather than the workload of the individual judges. By **the workload of the court**, I mean nothing more than **the cumulative workload of the members of the court**. This has the consequence that the court's workload is greater when one case is heard by nine judges than it is when the same case is heard by five judges. Viewed in one light, this is a mistake: at the margin, a five-judge case is no more onerous than a seven judge. Yet viewed from a logistical perspective, this makes sense: the workload of the court bites more when nine judges are occupied with one case, leaving too few to form a second panel.

Since workload is always measured at a point in time, I use information on the court's total workload six weeks before the first scheduled date of hearing.

The discussion of the court's own criteria for larger panel sizes highlighted several factors, of which I discuss three: case importance; constitutional issues; and human rights claims. I thereby ignore two other issues which, though I am sure they are important, are only intermittently so (departure from previous precedent, and conflict with decisions of the JCPC).[10]

As far as **case importance** is concerned, I use a measure based on the

[10] I can find only five cases in which the court considered whether or not it should depart from precedent, and in one of these cases (*Jones v Kernott*, [2011] UKSC 53 at 2) the court raised the issue itself.

number of generalist law reports in which the case below is reported.

The Supreme Court exists to decide points of law of general public importance. Yet not all of the cases which the Supreme Court is asked to hear are of such importance. The case of *Oliver v Symons* ([2012] EWCA Civ 267) is a good example. The Olivers of County Durham were annoyed that their neighbours had put a gate across a piece of farmland, a gate that the Olivers thought interfered with their right of way. They did not strictly need this right of way to get access to their farmland, since a new road had been built which would have allowed them to reach their farmland, but they were worried that the new road would get muddy if they drove over it in their farm vehicles.

Lord Justice Elias, who wrote the lead opinion in the case, thought that the case "ought never to have come near a court", and that the considerable costs incurred by the parties were "wholly out of proportion to the practical importance of the issue". He dismissed the Olivers' appeal, but this did not dissuade the Olivers, who sought permission to appeal to the Supreme Court. Unsurprisingly, the Supreme Court refused permission, finding that the case involved "no point of law of general public importance".

Lord Justice Elias and three judges of the Supreme Court agreed that *Oliver v Symons* was an unimportant case: indeed, it sometimes seems as if the only ones who did not agree were the Olivers and those representing them. But judges of the Supreme Court do not just form binary judgments about whether cases are "important" or "unimportant"; they also make finer distinctions between cases of greater and lesser importance. This is made explicit in the court's rules regarding the size of panels,[11] which mention cases "of *high* constitutional importance" or "of *great* public importance" (my emphasis). It is implicit in judges' behaviour during oral argument and in the writing of opinions. Since judges make distinctions of degree regarding the importance of each case, and

[11] https://www.supremecourt.uk/procedures/panel-numbers-criteria.html

since these distinctions affect their behaviour, any attempt at modelling judges' behaviour must employ some measure of the relative importance of each case.

I propose just such a measure below, and when I use it I will normally talk about importance *simpliciter*. But case importance — or salience, as political scientists often call it — can be viewed from two different aspects (Clark, Lax, and Rice 2016) which it is important to discuss before proceeding to measurement.

The first of these is a temporal aspect: cases can be important at the time they are decided, or they can be important in retrospect (Epstein and Segal 2000, 67). Contemporaneous and retrospective importance do not always go hand in hand. Some cases grow in importance over time. A good example is *Central London Property Trust Ltd v High Trees House Ltd* ([1947] KB 130), a case which was judged by subscribers to the Incorporated Council of Law Reporting (ICLR) to be one of the fifteen most important cases between 1865 and 2015, but which was described by a contemporary law report as seeming "scarcely reportable, much less epoch-making" (Neuberger 2015, 11; Magrath 2000).

It is the research question which determines whether one should measure contemporaneous or retrospective importance. Research into whether courts overturn "important" precedents at lower rates (Brenner and Spaeth 1995), or research evaluating judges on the basis of the cases they have decided (Maitra and Smyth 2004) would normally be expected to use a measure of retrospective importance. An explanation of judges' behaviour, on the other hand, must use a measure of contemporaneous importance, as I do here.

The second aspect of case importance is the substantive aspect. The substantive aspect asks: in virtue of what characteristics is the case important? Is it important because it is legally important, because it is politically important,

or because it is economically and financially important? These different understandings of substantive importance can diverge. Although many politically important cases will also be legally important, many legally important cases do not have the slightest political relevance.

Just as with the temporal aspect of important, it is the research question which determines the measure used. The measure I set out below is a measure of legal importance. Modern British judges are lawyers, not politicians:[12] what matters to lawyers matters to them. This may seem so obvious as to not require saying. I mention it because research on SCOTUS has typically used measures of importance which seem to track political importance, and, to the extent that it has distinguished between legal and political importance, has opted in favour of the latter (Cook 1993, 1131).

For example, the most widely used measure of the importance of cases decided by SCOTUS, developed by Lee Epstein and Jeffrey Segal, is based on whether or not the case was reported on the front page of the *New York Times* on the day following the judgment being handed down. Around 15% of cases are important, or salient, according to this measure. The measure has been criticised for being insufficiently "lawyerly": George and Guthrie (2008) state (p. 1844) quite baldly that "there is no reason to believe that a New York Times reporter can accurately distinguish important from unimportant cases."[13] I am sure that the measure succeeds in identifying cases which are important, both politically and legally: but I am also sure that the reason it does it because the legal and the political are so closely commingled in the United States, and not because newspapers of record place important legal developments on their front

[12]This has not always been the case. In the nineteenth century most judges had been politicians (Laski 1926).

[13]At the time the article was published, the *New York Times*' Supreme Court correspondent was Adam Liptak, who holds a JD from Yale and formerly advised the *New York Times* on legal matters. Liptak was preceded by Linda Greenhouse, who also has a post-graduate degree in law from Yale.

pages.

Other measures canvassed but ultimately rejected by Epstein and Segal do seem more lawyerly in nature, and more prone to identify cases of contemporaneous legal importance. One measure is whether or not the case was highlighted in *US Law Week*; this measure is rejected because of inconsistencies in the way it presented salient cases. A second measure is whether or not the case featured in the *Harvard Law Review*; but this too is rejected because of inconsistencies over time.

These rejected measures point the way to a measure of importance for cases heard by the Supreme Court, which both reflects the legal importance of the case and reflects the importance of the case at the time it was heard by the Supreme Court. I suggest that **the contemporaneous legal importance of a case heard by the Supreme Court can be measured by the number of generalist law reports which reported the case from which appeal lay**.

By a *generalist* law report, I mean a law report which covers cases of many different types within a given jurisdiction. I count four such law reports in England and Wales (the *Weekly Law Report*, the *All England Law Reports*, the *Solicitors' Journal Law Brief*, and the law reports published in the *Times* newspaper), three in Scotland (the *Session Cases*; the *Scots Law Times*; and *Green's Weekly Digest*), and one in Northern Ireland (*Northern Ireland Reports*).

Law reports have been used to measure the importance of cases in the House of Lords before (Blom-Cooper and Drewry 1972, 248). My measure is unusual in that I use reports of the case in the court *a quo*. I do this because I need a measure which is not just contemporaneous to the decisions of the judges of the Supreme Court, but which is prior to that behaviour. The Epstein and Segal measure of case importance discussed above is contemporaneous: front page

reports on SCOTUS decisions do not become available only after the passage of several years. It is not, however, prior to the decision: SCOTUS must decide something before the New York Times can report on it.[14] Law reports of the case *a quo* do not suffer from this problem.

Law reports are a good measure of the importance of cases because law reports in England and Wales have long been characterised by a broadly consistent approach to identifying important cases.[15] The generalist law journals routinely describe their role as one of identifying the most important cases. If pressed for details on the meaning of importance, some genuflection is usually made in the direction of the principles set out by Nathaniel Lindley in 1863: important cases are cases which

- "introduce, or appear to introduce, a new principle or a new rule";
- "materially modify an existing principle or rule";
- "settle, or materially tend to settle, a question upon which the law is doubtful";
- or cases which "for any reason are peculiarly instructive" (Magrath 2000, xi)

Cases reported in a single generalist law journal are thus more likely to be important. Since reasonable reporters may disagree about the importance of each case, cases reported in multiple generalist law journals are more likely still to be important. Indeed, some degree of disagreement between law reports is necessary for the measure to be an ordinal measure (the case is more or less important) rather than a dichotomy (the case was reported by all or reported

[14]Researchers may, of course, look at newspaper coverage of stages earlier in the process, such as reports of oral argument. See Clark, Lax, and Rice (2016) for a discussion of the considerable labour this effort entails.

[15]This also applies to Scotland and Northern Ireland. The Scottish Council of Law Reporting describes its criteria in a fashion almost identical to that described below. See "The Law of Judges: Precedent and the Criteria for the Reporting of Cases", `https://www.youtube.com/watch?v=E8MnKD5q1WE`

by none).

The measure treats Scotland and Northern Ireland differently. The maximum score for an English case is five, whereas the maximum for a Scots case is four (if the *Times* lends the Scots law journals a hand), and the maximum for a Northern Irish case is just one. The question then is whether or not these differences in the maximum possible values reflect true differences in the relative importance of Scots and Northern Irish law compared to English law.

This is not a question which admits a ready, empirical answer. There is no common yardstick against which cases from these different jurisdictions may be set. Though some appeals from different parts of the United Kindom are sometimes heard together, these conjoined appeals are too few in number and too exceptional in their subject matter to bridge the different jurisdictions.[16] Nor does any series of generalist reports straddle the different jurisdictions adequately. In matters of law, the *Times* remains an English and Welsh newspaper: it has reported on just 3 Scottish cases which were subsequently heard by the Supreme Court, and 0 from Northern Ireland. The measure thus rests on the political economy of the market for law reports: if the Scottish legal system can sustain three times as many generalist law reports as the Northern Irish legal system, it is (up to) three times more important.

One answer to this problem of commensurability is to *dilute* the problem. Rather than compare two well-defined sets of generalist journals, we could evaluate cases on the basis of a far larger (but less well-defined) set of specialist reports, some of which (due to the type of law, whether it be human rights law, European Union law, or something else) have the happy property of spanning jurisdictions. Since there are many more specialist journal than there are generalist journals, this would also result in a finer-grained measure of importance.

[16] Cases which involved appeals from Scotland and England: [2013] UKSC 63. Cases which involved appeals from Northern Ireland and England: [2011] UKSC 18, [2013] UKSC 61.

Unfortunately, this measure faces a different problem. The problem is not, as might be thought, one of different thresholds. It is true that reporters "who specialise in [one] department are quick to diagnose the extension of, or exception to, a principle" (Moran 1948, 75), but this difference of threshold would still allow measures of relative importance, in the same way that consistently combining measurements of temperature in degrees centrigrade and degrees Kelvin would still allow measures of relative heat. Rather, the problem concerns the relationship between the area of law and the number of specialist reports it can sustain.

Criminal law, for example, sustains fewer specialist journals than commercial law. This may be because criminal law is less important, or because commercial law is more variegated, or because impecunious criminal lawyers cannot afford to subscribe to a large number of specialist reports. Of these different explanations, only the first can dismissed. Across different areas of law, cases which have equal importance as measured by the number of *generalist* law reports have systematically different levels of importance as measured by the numbers of *specialist* law reports. A case from England and Wales which is reported in just one of the generalist journals can be expected to reported in 3.5 specialist reports if it is a case which deals with Chancery law, 3.4 if it deals with Family law, but only 1.4 if it deals with criminal law.[17]

We are therefore pushed back to using the count of generalist reports as a measure of case importance. Figure 3.1 shows the distribution of these scores for all 506 cases heard by the Supreme Court which were not the result of a reference question. Though the largest single category of cases is cases which were not reported at all, most cases were reported by at least one generalist law

[17]This claim is based on a linear regression on English and Welsh cases heard by the UK Supreme Court, where the dependent variable is the count in specialist reports (defined as any report which is not generalist), and where the independent variables are the area of law (as described above), the count in generalist reports, and the interaction between the two.

report, but not by very many more: the average across all cases is just 1.14.

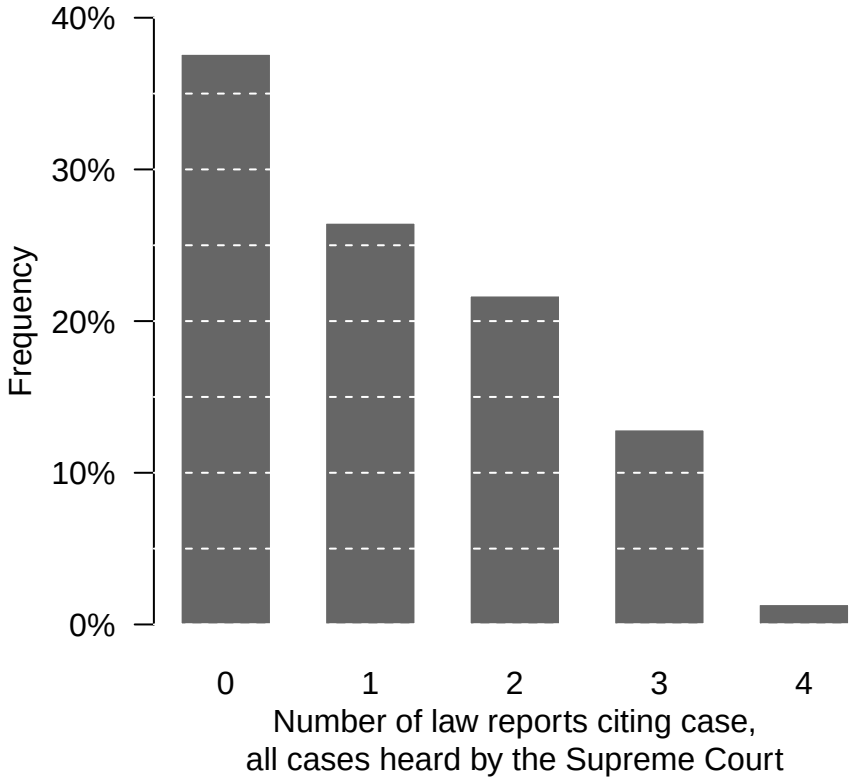

Figure 2: Importance of cases heard by the Supreme Court

To some extent the validity of this measure of contemporaneous legal importance depends on the results that follow. If we were to find, using this measure, that more important cases were *not* more likely to be heard in larger panels, we would be more likely to find fault with the measure than with the Supreme Court's decision-making.[18]

[18]This might suggest (per the Duhem-Quine thesis: Stanford (2013)) that the claim "the importance of a case is a factor in the Supreme Court's decision-making" is unfalsifiable, for

A measure of **constitutional importance** would require some understanding of what the uncodified British constitution requires; I cheerfully pass on this task, and fall back upon a dichotomous variable which measures whether or not a devolution issue was raised in the case, as it was in 3% of cases. I include a further dichotomous variable which has a value of one if the case referred to a specific article of the European Convention on Human Rights (something which featured in 25% of cases).[19]

The final account of panel size, based on avoiding non-compliance, introduces two further and final variables: **government participation**, and the **risk of government non-compliance**. To track government participation, I use a dichotomous variable which has the value one where central government acted as the appellant or the respondent. I exclude cases in which governments of other countries appeared (such as *Norris v. Government of the United States of America*, [2010] UKSC 9), cases in which central government appeared as intervenor, and cases in which non-ministerial departments or other non-departmental public bodies appeared. By this measure, central government appeared in 31% of cases.

To measure the risk of government non-compliance, I once again base my measure on the measure of opinion below. Where the government is the appellant, and where opinion below is more favourable to the appellant, the risk to the government of an unfavourable decision is low. Conversely, if opinion below is more favourable to the respondent, the risk of an unfavourable decision is higher. If all unfavourable decisions are, on average, equally objectionable to

---

any purported falsification can be explained away by rejecting the auxiliary assumption that the proposed measure of importance does what it claims to. Whilst this might be true in any single instance, this book tests the role of importance in multiple aspects of the Supreme Court's decision-making, and it need not be the case that case important is a factor in all of these aspects.

[19]This variable does not track whether or not the reference to ECHR rights was central to the case or not. Additionally, since it is based on the court's own judgment, it may exclude cases in which counsel raised human rights arguments which the court then ignored.

the government, and equally worth risking public backlash over, then we may assume that the greater the risk of an unfavourable decision, the greater the risk of non-compliance.

This means that in cases where the government is the respondent, we may use the measure of opinion below as a measure of the risk of government non-compliance. In essence, the risk of government non-compliance is equal to the (adjusted) proportion of judges who have found against the government. In cases where the government is the appellant, we have to reverse the measure so that it no longer measures support for the appellant, but support for the respondent.

Modelling

In order to model the choice of panel size, I use an ordered regression model. I do so because the outcome (the size of the panel) is not continuous (the panel may not have an even number of judges), but is ordered (five judges is less than seven judges, which in turn is less than nine judges).

Although the outcome is not continuous, it is often helpful to think of ordered regression models in terms of some underlying latent trait. In this example, we may think of cases as forming a continuum, from cases which are less deserving of a larger panel, to cases which are very much more deserving. Thinking in these terms simplifies the task of modelling, because the level of "deservingness" of each case can be modelled as a linear combination of variables and associated coefficients. This latent trait can then be turned into a specific outcome through a series of thresholds. Thus, a case merits a seven judge panel if its level of deservingness is larger than the threshold dividing five from seven judge panels, but smaller than the threshold dividing seven from nine judge panels.

The many different types of ordered regression models differ in how they

characterize the latent trait, and in their treatment of the thresholds. Here, I use an ordered probit regression model. The coefficients from this model are not easily interpretable, because they depend for their interpretation on the thresholds. Thus, although I report the results of the regression model, I spend more time on displays of specific effects.

Rather than run a single regression model, I estimate two models: one model which includes dummy variables for each of the several legal areas (Chancery law, Civil law, Family law, Northern Irish law, Scots law, and Administrative law (the reference category)). The inclusion of these additional controls causes our conclusions regarding some of the main effects to change. All references to effects refer to effects estimated from this second model, which includes controls for legal area.

Results

If we set all of the explanatory variables at their mean values (for continuous variables) or at their modal value (for categorical variables), then the model would predict that a five-judge panel is by far the most likely outcome (76% likely), followed by a seven-judge panel (20%) and a nine-judge panel (5%). Unsurprisingly, these predictions mirror the distribution of actual panels, *modulo* some slight differences which are the result of rounding error: the percentages of cases heard by five-, seven- and nine-judge panels are 79%, 16%, and 4% respectively.

A model which manages to recapitulate observed percentages and does nothing more is not a very useful model. By varying in turn some of the explanatory variables, we can explore the effects of those variables on panel size:

- by increasing the court's workload from its average workload, to a workload in the top quartile, we lessen the probability of a seven- or nine-judge

Table 2: Ordered regression model of panel size

	No. of judges	
	nJudges	
	(1)	(2)
Importance	0.316***	0.452***
	(0.103)	(0.115)
Probability	−0.372	−0.432
	(5.201)	(5.339)
Probability squared	−7.143	−7.684
	(5.743)	(5.875)
Workload	−0.006**	−0.005*
	(0.003)	(0.003)
HRA claim	0.608**	0.393
	(0.254)	(0.275)
Devolution issue	−0.071	1.467**
	(0.580)	(0.748)
Non-compliance risk	0.689	0.520
	(0.494)	(0.560)
Chancery		−0.616
		(0.438)
Civil		−0.282
		(0.364)
Criminal		0.434
		(0.505)
Family		0.593
		(0.484)
Northern Ireland		−0.156
		(0.714)
Scots		−1.922***
		(0.576)
Importance by probability	0.644	1.427
	(2.526)	(2.667)
Importance by probability squared	0.763	1.176
	(2.695)	(2.814)
5 judges|7 judges	1.337***	1.271***
	(0.393)	(0.438)
7 judges|9 judges	3.139***	3.127***
	(0.437)	(0.479)
Observations	504	504
Note:	*p<0.1; **p<0.05; ***p<0.01	

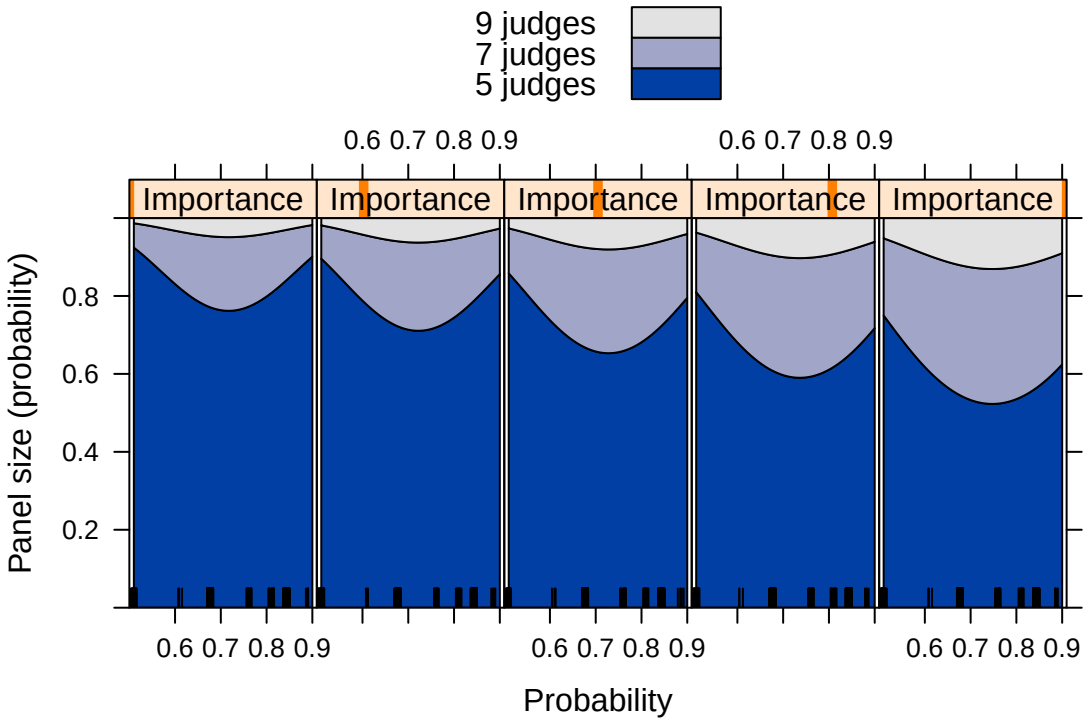

Figure 3: Probability of panels of different sizes, according to the importance of the case, and the probability of a correct decision

panels, but only very slightly. These probabilities fall to 18% (from 20%) and 4% (from 5%) respectively.

- introducing a human rights claim into the case raises the probability of a seven judge panel to 25%, and that of a nine judge panel to 7%, but this effect is only significant in the model which lacks controls for legal area.
- these effects are smaller than the effects of introducing a devolution issue, which raises the probabilities of seven- and nine-judge panels (to 40% and 18% respectively). Note however that this effect is entirely cancelled out in practice by the much lower panel sizes given to Scottish appeals.
- Finlly, when we move from a case which involves no risk of non-compliance to a case involving the average level of noncompliance, the probability of a seven judge panel increases to 22% (from 20%), and the probability of a nine judge panel increases to 6% (from 5%), but this effect is not significant.

These effects, though important, are secondary to the main theoretically-derived predictions regarding the probability of a correct decision and the importance of the case. Figure 2 shows the joint effects of these two variables.

The figure is split into five segments, which correspond to cases of lowest (the left-most segment) to the utmost importance (right-most segment). Within each segment, the horizontal axis runs from cases where the probability of a correct decision is very low, to cases where the probability of a correct decision is very high. Within each segment, the shaded areas show the probability of different panel sizes being appointed.

The probability of a five-judge panel (in dark blue) gradually decreases as we move from less to more important cases. Note, however, the way the dark blue shaded area contracts then expands within each segment. Very difficult cases (i.e., cases where the probability of a correct decision is very low) are

more likely to get a five-judge panel, as are very "easy" cases. This is exactly what was predicted by the model of optimal panel sizes.

Unfortunately for that model, the non-linear effects of the ex ante probability of a correct decision are not statistically significant. This is strongly suggested by the results reported in Table 1, but since the significance of an interaction term does not depend on the individual significance of its constituent terms, it is necessary to test a model with the square of probability to a model which omits it. That model (not shown here) has a very slightly worse fit to the data, but the difference in fit is not statistically significant.

At this stage, I am happy to wait for additional data to either confirm or disprove the counterintuitive suggestion that it is intermediate difficulty cases which are most likely to receive the largest panels. In my treatment of the jury theorem, I noted that the theorem is often interpreted as requiring two exclusive and exhaustive options: innocent or guilty, appeal allowed or appeal dismissed. In this paper, I ask for a more hesitant "not proven" verdict: although there is some evidence to suggest the jury theorem's insights hold, this evidence is not strong enough for conventional statistical significance filters. Yet nor is it so weak that we should reject it entirely.

Conclusions

In this paper, I set out three different accounts of the size of panels on the UK Supreme Court: one account which was built on Condorcet's Jury Theorem, and which suggested that the largest panels would be reserved for cases of intermediate difficulty; one account which was built on the statements of the court and of the judges who comprise it, and which suggested that the largest panels would be reserved for those cases raising human rights and constitutional issues, and possibly also for the most difficult cases; and one account which

suggested that the largest panels would be reserved for cases in which there was a risk that the court would have to rule against the government.

The first two accounts suggested, almost in passing, that more important cases would receive larger panels, and so it has proved. This provides some vindication of the measure of case importance set out above. The jury theorem also suggested that the court's workload might play a role, insofar as it heightens the opportunity cost of already-burdened judges participating in additional (large) panels. This variable also had a significant effect in the models above.

These findings regarding panel size are but half of the story. Future research must deal with the composition of the resulting panel. I have considered the issues of panel size and panel composition as separate and sequential decisions. As the anecdote about *Radmacher v. Granatino* suggests, however, these decisions are sometimes commingled.

References

Alarie, Benjamin, Andrew James Green, and Edward 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.

Birks, Peter. 2000. *English Private Law*. Oxford: Oxford University Press.

Blanes i Vidal, Jordi, and Clare Leaver. 2011. "Are Tenured Judges Insulated from Political Pressure?" *Journal of Public Economics* 95 (7). Elsevier: 570–86.

Blom-Cooper, Louis Jacques, and Gavin Drewry. 1972. *Final Appeal: A Study of the House of Lords in Its Judicial Capacity*. Oxford: Clarendon Press.

Brenner, Saul, and Harold J Spaeth. 1995. *Stare Indecisis: the Alteration of Precedent on the Supreme Court, 1946-1992*. Cambridge University Press.

Briant, Sophie. 2011. "Dialogue, Diplomacy and Defiance: Prisoners' Voting Rights at Home and in Strasbourg." *European Human Rights Law Review* 3: 234–52.

Buxton, Richard. 2009. "Sitting En Banc in the New Supreme Court." *Law Quarterly Review* 125 (April): 288–93.

Clark, Tom S, Jeffrey R Lax, and Douglas Rice. 2016. "Measuring Salience as a Latent Variable." *Law & Courts*, 8.

Condorcet, Nicolas de. 1785. *Essai Sur L'application de L'analyse La Prob-*

abilit Des Dcisions Rendues La Pluralit Des Voix. Paris: Imprimerie Royale.

Condorcet, Nicolas de, and Keith Michael Baker. 1976. *Condorcet: Selected Writings*. Edited by Keith Michael Baker. Indianapolis: Bobbs-Merrill, 1976.

Cook, Beverly B. 1993. "Measuring the Significance of US Supreme Court Decisions." *The Journal of Politics* 55 (04). Cambridge Univ Press: 1127–39.

Cooper, Davina. 1996. "Institutional Illegality and Disobedience: Local Government Narratives." *Oxford Journal of Legal Studies* 16 (2). JSTOR: 255–74.

Darbyshire, Penny. 2011. *Sitting in Judgment: the Working Lives of Judges*. Bloomsbury Publishing.

De Tocqueville, Alexis. 1835. *Democracy in America*.

Dickson, Bruce. 2007. "The Processing of Appeals in the House of Lords." *Law Quarterly Review* 123: 571–601.

Dworkin, Ronald. 1985. *A Matter of Principle*. Oxford: Oxford University Press.

Epstein, Lee, and Jeffrey A Segal. 2000. "Measuring Issue Salience." *American Journal of Political Science* 44 (1). JSTOR: 66–83.

George, Tracey E, and Chris Guthrie. 2008. "The Threes: Re-Imagining Supreme Court Decisionmaking." *Vand. L. Rev.* 61 (6). HeinOnline: 1825–60.

Gibson, James L, and Gregory A Caldeira. 1998. "Changes in the Legitimacy of the European Court of Justice: A Post-Maastricht Analysis." *British Journal of Political Science* 28 (01). Cambridge Univ Press: 63–91.

Knight, C.J.S. 2012. "The Supreme Court Gives Its Reasons." *Law Quarterly Review* 128 (October): 477–81.

Laski, Harold J. 1926. "The Technique of Judicial Appointment." *Michigan Law Review* 24 (6). JSTOR: 529–43.

List, Christian, and Robert E Goodin. 2001. "Epistemic Democracy: Gen-

eralizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9 (3). Wiley Online Library: 277–306.

Magrath, Paul. 2000. "Introduction." *The Law Reports* Special issue: ix–x.

Maitra, Pushkar, and Russell Smyth. 2004. "Judicial Independence, Judicial Promotion and the Enforcement of Legislative Wealth Transfers—an Empirical Study of the New Zealand High Court." *European Journal of Law and Economics* 17 (2). Springer: 209–35.

Ministry of Justice. 2015. "Judicial Salaries and Fees: 2015 to 2016." `https://www.gov.uk/government/publications/judicial-salaries-and-fees-2015-to-2016`.

Moran, Clarence Gabriel. 1948. *The Heralds of the Law*. Stevens.

Neuberger, David. 2015. "Reflections on the ICLR Top Fifteen Cases: A Talk to Commemorate the ICLR's 150th Anniversary." `https://www.supremecourt.uk/docs/speech-151006.pdf`.

Paterson, Alan. 2013. *Final Judgment: the Last Law Lords and the Supreme Court*. Bloomsbury Publishing.

Stanford, Kyle. 2013. "Underdetermination of Scientific Theory." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2013. `http://plato.stanford.edu/archives/win2013/entries/scientific-underdetermination/`.

Staton, Jeffrey K. 2006. "Constitutional Review and the Selective Promotion of Case Results." *American Journal of Political Science* 50 (1): 98–112.

Stockley, Nicola, ed. 2012. "The Highest Court in the Land: Justice Makers." BBC 4.

Stone, Alec. 1992. *The Birth of Judicial Politics in France: The Constitutional Council in Comparative Perspective*. Oxford: Oxford University Press.

Vanberg, Georg. 2004. *The Politics of Constitutional Review in Germany*. Cambridge: Cambridge University Press.