

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press  
32 Avenue of the Americas, New York, NY 10013-2473, USA  
www.cambridge.org  
Information on this title: www.cambridge.org/9780521680035

© Cambridge University Press 2009

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2009

Printed in the United States of America

*A catalog record for this publication is available from the British Library.*

*Library of Congress Cataloging in Publication data*

A quantitative tour of the social sciences / [edited by] Andrew Gelman, Jeronimo Cortina.  
p. cm.

Includes bibliographical references and index.

ISBN 978-0-521-86198-4 (hardback) – ISBN 978-0-521-68003-5 (pbk.)

I. Social sciences – Methodology. 2. Social sciences. I. Gelman, Andrew.

II. Cortina, Jeronimo. III. Title.

H62.Q365 2009

300.72–dc22 2008054986

ISBN 978-0-521-86198-4 hardback  
ISBN 978-0-521-68003-5 paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables, and other factual information given in this work are correct at the time of first printing, but Cambridge University Press does not guarantee the accuracy of such information thereafter.

# A QUANTITATIVE TOUR OF THE SOCIAL SCIENCES

Edited by

ANDREW GELMAN

*Columbia University*

JERONIMO CORTINA

*University of Houston*

 **CAMBRIDGE**  
UNIVERSITY PRESS

## 17. Modeling Strategy in Congressional Hearings

In the previous chapter, we concluded that manipulating the media environment is a central concern for strategic players in the “game” of Supreme Court nominations. From this perspective, the president and his allies must try to establish a favorable media environment for the nominee. Individuals opposed to the nominee and the president must do the reverse: establish a hostile environment. Of course, a favorable environment may include flying under the radar – no attention in the media, or very little.

How can strategic actors manipulate the information environment? A useful distinction is between direct and indirect manipulation. By “direct manipulation,” I mean, for example, the president going public (making speeches) on behalf of the nominee, interest groups staging protests and demonstrations, and members of the Judiciary Committee shaking their finger at the nominee and asking “gotcha” questions about judicial doctrine. These are direct manipulations because the media may deem them newsworthy and broadcast them or report on them.

Indirect manipulation is somewhat more subtle. The idea is that you engage in actions that make direct manipulations (your own or others) more or less effective, or encourage reporters to write the right kinds of stories on their own. Consider, for example, the president’s choice of a moderate versus an extreme nominee or a highly qualified one versus a marginally qualified one. These choices are indirect manipulations of the media because they affect how the media portray the nominee. These choices also make the president’s statements about the nominee more or less credible (at least arguably). As we have seen, selection of a moderate, able nominee will encourage favorable press reports, while selection of an extremist hack will tend to generate negative publicity. As another example of indirect manipulation, an interest group compiles incriminating statements – for example, “The Book of Bork” assembled by anti-Bork activists – and turns them over to the press. Thus, the interest group does the work of investigating the nominee for the press, which leads to increased press coverage of the kind the group wants.

What about the Judiciary Committee? How can the chairman or the majority party of the committee indirectly manipulate the press? Suppose the chair is hostile to the nominee. Then an obvious way to indirectly manipulate the media is to do a thorough investigation and uncover some really damaging evidence of improper behavior. (We saw in the previous chapter the huge impact of scandals

on media coverage.) Then, during the hearing itself, the chair and like-minded members of the committee can beat the nominee over the head with his or her transgressions. Thus, the indirect manipulation – searching for a scandal – enhances the direct manipulation, the public hearings.

One might suppose that a friendly chair should do the same: investigate thoroughly to reveal how clean the nominee is, then publicize it during the hearings. But, as we saw, when it comes to Supreme Court nominations, no news is good news: Absent a scandal and a broadly hostile Senate, a nominee is practically sure to be confirmed. So, a better bet for a friendly chair is to move the nominee through the process so as to avoid accusations of scandal. In practice, this means, hustling the nominee through as fast as possible.

In this chapter, I explore these ideas in more detail. We’ll take a close look at the duration and intensity of public hearings during Supreme Court nominations. I’ll try to convince you that the public hearings are longer and more intense when (1) the chair of the Judiciary Committee is hostile to the nominee, (2) the chair has some red meat to work with, that is, when the prior search or good luck uncovered a scandal or other bad conduct, and (3) staff resources are more abundant, so prolonged hearings are less costly of effort to senators. We’ll also find some evidence suggesting that hearings became longer and more intense beginning in the late 1960s in the wake of the controversial Warren Court.

You may be willing to believe all this already, without looking at any data or models. So, at some level, the findings may not be terribly surprising. But my point isn’t to unfold a shocking revelation about American politics, though I do want to explore this material carefully. Instead, I want to return to a theme I introduced in Chapter 15: the interplay of theory and data. I will show you, in a simple way, how this works.

The analysis proceeds in six steps. First, we’ll probe the logic of the situation. To do so, we’ll develop a simple formal – that is, mathematical – model of the chair’s behavior. (I claim that the model is a more general representation of congressional investigations of the executive.) We’ll interrogate the model deductively to generate specific empirical predictions. Thus, we’ll know in a very precise way how a set of theoretical assumptions maps onto empirical hypotheses.

Second, we’ll try to measure the key variables indicated by the theory. This will involve a fair amount of work (all too common, I’m afraid). Third, we’ll search the data for structure, looking for the predicted patterns but also for unexpected ones. As before, we’ll rely on artful visualizations and highly flexible, nonparametric modeling.

Fourth, we’ll fit parametric models to the data, testing the formal model’s predictions. We’ll use models appropriate for count data and briefly discuss some issues that arise there. Fifth, we’ll ask what could go wrong with the analysis – how much should I believe the theory?

Finally, we’ll address some normative questions. Given what we’ve found, how should we evaluate the kind of partisan behavior emphasized by the theory?

### A NEW THEORY OF CONGRESSIONAL HEARINGS

What do political scientists know about the length and intensity of congressional hearings? Is there a standard model for the process? In fact, there is an empirical

literature that examines how frequently Congress holds hearings and what the hearings are about (Aberbach 1991, 2002; Baumgartner and Jones 2002). There is a small theoretical literature that treats hearings as signals from a congressional committee to an agency about the committee's interest in a matter (Cameron and Rosendorff 1993; Ferejohn and Shipan 1989) or as a signal from a committee to floor members of Congress about the committee's knowledge of an issue (Diermeier and Feddersen 2000). None of the theoretical works focus on hearings as an opportunity for congressmembers to gain publicity or to damage the investigation target. So, let's build a new model of public hearings from scratch, bearing in mind the case of Supreme Court nominees. Our theory will be based on the idea of publicity-seeking partisans who use hearings to generate news stories.

We'll begin by assuming that (1) public hearings generate negative messages about a nominee, (2) negative messages bring a reward to the chair of the committee, at least when he or she is ideologically estranged from the nominee, and (3) public hearings are costly to the chair to conduct (time and effort costs and forgone leisure). We will assume that the objective of the chair is to maximize his or her net rewards, that is, maximize rewards less effort costs. The pieces of our model are negative messages, a production function for creating negative messages, a reward function from generating negative messages, and a cost function for producing negative messages. We'll think a little harder about each of these pieces, then put them all together and consider the chair's behavior in the model.

Let's denote the number of negative messages as  $n$  and the negative message production function as  $n = f(x; \theta)$ , where  $x$  indicates the duration of the hearing and  $\theta$  the stock of "bad stuff" known about the nominee. We'll take  $\theta$  as fixed; it reflects the prior work by the committee staff and interest groups in digging up dirt on the nominee. The chair builds negative messages by extending the public hearings. So, negative messages increase with the length of the hearings. A seemingly sensible assumption is declining marginal productivity; longer hearings mean more bad messages, but at a declining rate from each additional day of hearings. Another sensible assumption is that the more ammunition the chair has to use against the nominee (the bigger is  $\theta$ , the greater the number of bad messages for the same level of effort. In addition – and this is critical – we assume "increasing differences" in duration and dirt. A figure is useful at this point, especially to explain increasing differences.

Examine Figure 17.1, which illustrates the type of production function I have in mind. The lower curve shows the effect of longer hearings ( $x$ ) at a lower level of dirt ( $\theta$ ), the higher curve the effect at a higher level. You can see the declining marginal effect of hearing duration with each curve, in that they tend to flatten out. You can also see increasing differences at work: As duration increases, the difference between the high-dirt curve and the low-dirt curve increases. At any duration, the marginal effect of duration is greater when  $\theta$  is greater.<sup>1</sup> An example of a function like this is  $n = \theta \log(x)$ , which is the parametric form of the curves in the graph.

<sup>1</sup> If the production functions are differentiable, the increasing differences assumption is just  $\frac{\partial^2 f}{\partial x \partial \theta} > 0$ : More dirt raises the marginal increase in negative messages from longer hearings.

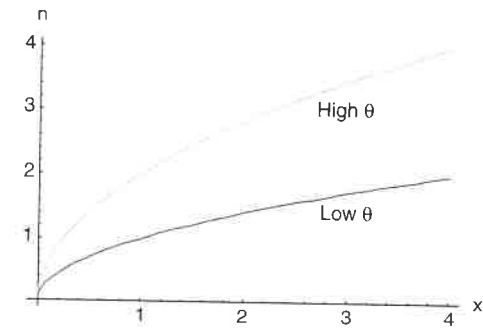


Figure 17.1. Hypothesized production functions for negative messages.

Now let's consider the return to the chair from the production of negative messages. We'll assume that the chair's return is a function of negative messages – and that the effect of negative messages depends on the ideological distance between the chair and the nominee. Let's call this ideological distance  $\delta$ . Then we can denote the return to the chair from negative messages as  $r(n; \delta)$ . How do we imagine this return function works? If the chair is close enough to the nominee, he or she doesn't want any negative messages about the nominee:  $r(n; \delta)$  is negative for positive  $n$  if  $\delta$  is small enough. But beyond a certain degree of ideological estrangement,  $r(n; \delta)$  will be positive for positive  $n$ . Plausibly, the greater the ideological distance between the chair and the nominee,  $\delta$ , the more the value to the chair of negative messages.

Again, it may be helpful to look at a picture, so examine Figure 17.2, which illustrates a return function. The lower curve is the return from negative messages to a chair who is ideologically proximate to the nominee. As shown, the return to the chair is negative, so he or she wants no negative messages about the nominee. The middle line is the return function to a chair who is somewhat estranged from the nominee, say, a moderate Democrat facing a somewhat conservative Republican nominee. The top curve is the return function for a very liberal chair who faces a very conservative nominee. An example of a return function like this is  $r(n; \delta) = (\delta - \bar{\delta})n$ , where  $\bar{\delta}$  represents the critical distance at which the return function switches from negative to positive.

Our third building block is the cost of holding public hearings. Here we'll assume a cost function  $c(x; \kappa)$  in which costs rise with the duration of the hearing but decline with, say, the abundance of staff resources  $\kappa$ . Presumably as well, the cost of zero hearings is zero, that is,  $c(0) = 0$ . An example of such a function is  $x/\kappa$ .

We now have our building blocks. Let's assemble them to describe the chair's objective. Pursuing the idea that the chair is interested in the net returns from hearings, we have

$$\Pi(x, \theta, \delta, \kappa) = r(f(x, \theta), \delta) - c(x, \kappa). \quad (1)$$

This is just the return function minus the cost function. We imagine that the chair wishes to maximize this function by choosing a hearing duration  $x$ . In Figure 17.3, I've used the component equations I mentioned previously to construct a specific net

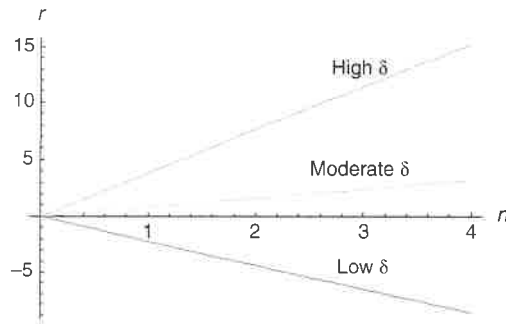


Figure 17.2. Hypothesized return functions from negative messages.

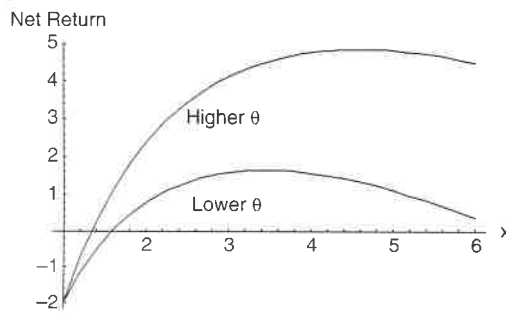


Figure 17.3. Hypothesized net returns from hearing lengths at high and low levels of dirt.

return function at two different levels of  $\theta$ . As you can see, at least in this example, there is a clear optimal length for each level of dirt. Call this length  $x^*$ .

An obvious question is, when will equation (1) have a well-defined maximum? This is not a trivial matter and has been the subject of much thought by mathematicians. You can probably see that if the net return function is continuous and the allowable hearing duration is a continuous closed interval, like 0–6, the net return function must have a maximum value in the interval. In such a case, if the net return function is differentiable, you can find the (or at least a) maximum using simple calculus. Things can get more complicated, though, if, for example, the net return function has jumps or the allowed duration of hearings takes only integer values, and so on.

At this point, we could go off on a long digression on the theory of optimization, but instead, let's keep our eye on the theory–empirics link. To do so, we need to think harder about the optimal hearing duration  $x^*$ . Presumably  $x^*$  is the length of public hearings the chair will choose; at least, that's our assumption. We can observe hearing lengths across different nominations, which will become the outcome variable of interest. Now the optimal hearing length will be a function of the exogenous variables in equation (1), specifically, the amount of dirt  $\theta$ , the ideological distance  $\delta$  between the chair and the nominee, and the size of staff resources  $\kappa$ . In other words, optimal length is itself a function  $x^*(\theta, \delta, \kappa)$ . Over a

series of nominations, we may be able to observe the values of these variables, which will become our predictor variables. So, if we can coax our theoretical model into making predictions about the behavior of the function  $x^*(\theta, \delta, \kappa)$  as  $\theta$ ,  $\delta$ , and  $\kappa$  change, we will have some testable predictions relating our outcome and predictor variables.

Let me introduce some jargon. Assuming that  $x^*(\theta, \delta, \kappa)$  is differentiable (which we don't actually have to assume), the predictions we want concern  $\frac{\partial x^*}{\partial \theta}$  (the change in  $x^*$  as  $\theta$  increases),  $\frac{\partial x^*}{\partial \delta}$  (the change in  $x^*$  as  $\delta$  increases), and  $\frac{\partial x^*}{\partial \kappa}$  (the change in  $x^*$  as  $\kappa$  increases). These three partial derivatives are called “comparative static multipliers,” which is probably not completely transparent terminology. But the terminology is too well established to change, so let's go ahead and use it.

From a theoretical point of view, the comparative static multipliers summarize the model's empirical content – for most practical purposes, they are its empirical content. Rebecca Morton, in her useful book *Methods and Models* (1999), discusses some other ways to test theoretical models in the social sciences. But in general, comparative statics is the starting point for empirical understanding of formal mathematical models in political science.

So, our problem is to derive the comparative static multipliers for our model. I should note that, typically, we will not be able to establish exact numerical values for the comparative static multipliers. For example, we will not be able to say, “The theory predicts that if the stock of bad news about a nominee goes up by one unit, the hearing duration will increase by five hours.” Predictions like this demand too much from theory. But we may be able to get predictions about the sign of the multipliers – for example, whether  $x^*$  gets larger or smaller or stays the same (or we can't tell) as the stock of bad news about a nominee increases.

How do we derive the comparative static multipliers in our model? Basically, there are three ways to do it. The first is to derive a closed-form solution. This is the easiest method to understand. The idea is to assume specific functional forms for the production function, return function, and cost function and then solve for the function  $x^*(\theta, \delta, \kappa)$ . If we are careful about specifying the underlying functions, this may be possible. Then, given an explicit  $x^*(\theta, \delta, \kappa)$  function, we can see how  $x^*$  changes as the exogenous variables change, either through inspection, calculus techniques, or direct calculation.

I will work out a closed-form solution next, but the obvious difficulty with this approach is that the results are hostage to the specific equations we assume. Consequently, there may be little or no generality to the results. Perhaps if you used a somewhat different but equally plausible return function (for instance), you might get completely different predictions. And any test of the model based on a closed-form solution will actually be a test of the full set of assumptions – not just the general equation (1) but all the specific forms for each of the components.

A second method is to use some powerful results in the theory of robust comparative statics. In recent years, economic theorists have worked out general results concerning the behavior of equations like (1).<sup>2</sup> If you feel comfortable

<sup>2</sup> For example, see Milgrom and Shannon (1994). An accessible introduction written for political scientists is Ashworth and Bueno de Mesquita (2006); also see Athey, Milgrom, and Roberts (1998) for the perspective from economics.

making some rather broad assumptions about the components of equation (1) – the kinds of assumptions we made earlier, like “increases in  $x$  and  $\theta$  with increasing differences between  $x$  and  $\theta$ ” – you can apply powerful theorems directly to the problem and immediately get the comparative static results. I was careful to make the right kinds of assumptions as we went along, so we can do this. I won’t go into too much detail about this, but a nice implication of this approach is that, because we know that the results are perfectly general so long as we maintain the underlying assumptions about the component equations, we can use any tractable functional forms we want and derive closed-form solutions whose behavior is (in a qualitative sense) perfectly general. These solutions, in turn, can supply specific estimating equations. (You may want to think about the logic of this for a minute.) So, combining the two methods is very attractive when taking theory to data.

A third method is somewhat intermediate between the specificity of the first approach and the generality of the second. Here you make some general assumptions about the shapes of the component equations but not their specific form, and also some strong assumptions about their smoothness and continuity. Then you can use calculus-based methods to work out the comparative statics. This approach typically involves quite a bit of algebra, so we won’t do any more with it.

Let’s assume the specific functional forms I used earlier, since we know that the results we’ll find in this case are actually quite general. Thus, equation (1) becomes

$$(\delta - \bar{\delta})\theta \log(x) - x/\kappa. \quad (2)$$

This function is what is shown in Figure 17.3 for the values  $\delta=3$ ,  $\bar{\delta} = .7$ ,  $\kappa = 1/2$ ,  $\theta=3$ , and  $\theta=4$ . (I picked these values through trial and error to get pretty shapes for the curves.) As you can see, there is a clear optimal hearing length, the point at which the net return is highest. This optimum length increases as the amount of available dirt increases. Make sure that you understand this point: The optimal length (the point where the curve is highest) shifts to the right as  $\theta$  increases. Simply from inspection of the curves, we can see our first comparative static in action.

I could redo the picture, varying each of the parameters in equation (2) to give a pictorial version of the comparative statics. Instead, let’s derive an expression for the optimum length. Using a little algebra,<sup>3</sup> I find this to be

$$x^*(\delta, \bar{\delta}, \theta, \kappa) = (\delta - \bar{\delta})\kappa\theta \quad (3)$$

This is actually pretty neat. The optimal length is simply the product of the three relevant parameters (counting  $(\delta - \bar{\delta})$  as a single parameter).

A point that jumps out of equation (3) is that if the chair is close enough to the nominee (if  $\delta \leq \bar{\delta}$ ), the chair will not want to hold any hearings at all:  $x^*$  is zero (obviously, hearing length can’t be negative). Or, if zero hearings aren’t feasible, the chair will hold the shortest feasible hearings. Before the late 1940s or so, the

<sup>3</sup> Take the first derivative of (2) with respect to  $x$  and set this equal to zero (i.e., assume an interior maximum), then solve for  $x$  at the interior maximum.

Senate Judiciary Committee sometimes allowed Supreme Court nominations to go to floor votes without holding public hearings. After the late 1940s, the committee stopped doing this; apparently, having no public hearings at all seemed unacceptably cavalier. But even then, some hearings were very brief, pro forma affairs. We’ll look at some data on this shortly.

Using equation (3), we can get our comparative static predictions directly.<sup>4</sup> When  $\delta > \bar{\delta}$  and  $\theta, \kappa > 0$ :

$$\begin{aligned} \frac{\partial}{\partial \delta} x^*(.) &= \kappa\theta > 0 \\ \frac{\partial}{\partial \theta} x^*(.) &= (\delta - \bar{\delta})\kappa > 0 \\ \frac{\partial}{\partial \kappa} x^*(.) &= (\delta - \bar{\delta})\theta > 0. \end{aligned} \quad (4)$$

In words, when the chair and the nominee are estranged ideologically, hearing duration lengthens as the degree of estrangement deepens, the stock of dirt increases, and the committee staff swells. If we just used the robust approach, we would get similar results, except that in general, each of the equalities would be weak (greater than or zero rather than greater than). Note that if  $\delta \leq \bar{\delta}$  (the nominee is close to the chair), these equalities all become (weakly) negative, implying a minimal hearing length.

We’ll take the model to data, proceeding in several steps. First, we’ll visualize the relationships nonparametrically, using extremely flexible functional forms. This is a good way to listen to the data – we’ll let the data talk to us without forcing them into any procrustean beds. This approach fits neatly with the robust comparative statics philosophy: We eschew specific functional forms and look for the broad patterns. But we’ll also estimate parametric equations inspired by our closed-form solution. This is attractive since equation (3) is so simple, and the parametrics allow us to do standard diagnostics on this regression – for example, making sure that the results are robust to influential outliers, and so on.

To summarize our progress so far: We have developed a new theory of congressional hearings, with an eye toward Supreme Court nominations. Our theory is based on the idea that congressmembers are ideological partisans. According to the theory, the chair of the Judiciary Committee drags out the hearings in order to publicize nasty bits of information about an ideological enemy. In contrast, the chair hustles an ideological soul mate through the hearings in order to limit any damage to him or her. Thus, public hearings on Supreme Court nominees are not a seminar in a philosophy class or law school. Rather, they are a kind of political warfare, at least when the chair and the nominee are ideologically estranged. Our key predictor is the ideological distance between the chair and the nominee. Also important is the availability of dirt on the nominee. Under some reasonable and fairly general assumptions, we worked out the empirical implications of our theory. Because we did this mathematically, we can be sure that our logic is tight.

Now let’s turn to the data.

<sup>4</sup> Just take the partial derivative of (3) with respect to the parameter of interest.

### DATA: DEFINITIONS AND OVERVIEW

The data come from all nominations from Hugo Black (1937) through Samuel Alito (2005). Two variables, length of the hearing (duration in days) and size of the hearing report (length in pages), represent the intensity of hearings and are our outcomes. Our theory suggests three predictor variables: (1) a measure of ideological distance between the nominee and the chair of the Judiciary Committee, (2) an indicator for scandal, measuring the dirt or bad news about the nominee at the time of the hearing, and (3) the number of Judiciary Committee staff. In addition, a little knowledge of history suggests a fourth: an indicator for senatorial courtesy (explained shortly).

#### Length of the Hearing

The first measure of intensity of public hearings is the number of days of hearings for each Supreme Court nominee, as reported in Rutkus and Bearden (2005). This measure is the actual number of days on which public hearings were held. For example, in the Powell nomination, the public hearing started on November 3, 1971, and ended on November 10, 1971, a total of eight days. However, hearings were not held on November 5, 6, and 7. Thus, the recorded number of hearing days is five rather than eight.

Using this measure, the mean duration of hearings was 3.3 days and the median was 2 days. The minimum length was zero days, the maximum 12 days (the Bork nomination), and the variance was a little over 10 days.

#### Volume of the Hearing Report

The Senate Judiciary Committee publishes an official report on each hearing. The report contains transcripts of the witnesses' testimony, senators' questions, and the nominee's answers, plus various items the senators want to put on the record. It seems natural to view the volume of the report as proportional to the intensity of the public hearing.

The average number of pages in the committee reports is 454, with a median of 128. The minimum length is zero and the maximum is 3,350 (again, the Bork nomination). The variance is huge.

Figure 17.4 shows the distribution of the number of pages in the reports, on the original and the log scale. At this point, it should be no surprise that the raw data do not follow a normal (bell-shaped) distribution. Taking the logarithm compresses the long right tail and hints that there may be two distinct populations.

#### Ideological Distance

The theory indicates that the more distant the nominee's ideology is from that of the Judiciary Committee chair, the longer the hearing will be (with all else unchanged). To estimate the distance from the nominee's ideology to the chair, we need to measure the ideologies of the chair and the nominee and put them on the same scale. This takes some work!

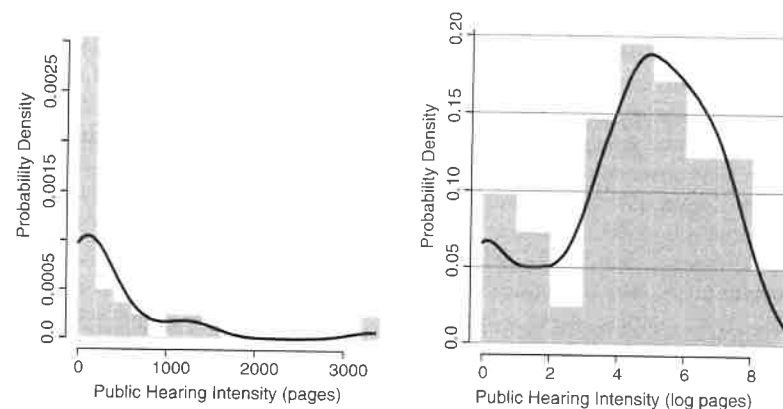


Figure 17.4. Distribution of intensity of public hearings for Senate nominees, as measured by number of pages in committee reports and shown on the log scale. Histograms show the data, and curves show density estimates. The histograms give a clearer picture of the data. The histograms would be improved by showing actual counts on the y-axes so that the reader could get a sense of the amount of data in each histogram bar.

Assigning a score to the Judiciary Committee chair is quite straightforward. We first identified the Judiciary Committee chairs at the time of each nomination and then located their DW-NOMINATE scores in Rosenthal and Poole's dataset.<sup>5</sup> I mentioned these scores in Chapter 15. Poole and Rosenthal's DW-NOMINATE scores, based on all recorded roll call votes in Congress, are widely used by political scientists as measures of congressmembers' ideologies and are generally considered highly reliable.

The process is much more complicated for the nominees and requires a little explaining. First, given our use of a roll call-based measure for senators, it might seem natural to do the same thing for Supreme Court justices. Then, given something like a DW-NOMINATE score for a justice, we could try to put the nominee measure and the senator measure in synch somehow. In fact, several scholars have tried to do this (Bailey and Chang 2001; Epstein et al. 2007).

I believe that voting scores for Supreme Court justices have their uses, but they are not well suited for studying nomination politics. First, there are no voting scores for failed nominees since they never made it to the Court (obviously). That creates a problem immediately. Second, even for the successful nominees, we are more interested in the way they were perceived at the time of the hearings than in how they actually voted later on the Court. The two are surely related, but they are not the same thing. Finally, there are good theoretical and historical reasons to believe that the voting behavior of Supreme Court justices may not be a simple expression of their preferences. For one thing, their votes surely reflect the changing docket of cases, which they themselves choose.

<sup>5</sup> Available at [www.voteview.com](http://www.voteview.com).



To see why this is a problem, consider the voting of an ideologically moderate justice when a conservative majority controls the docket. This justice, who will often split from the conservative majority and vote with liberals on the Court, is apt to look rather liberal. But if a bloc of liberals controlled the docket, this same moderate justice would often vote with conservatives, thus making him or her look rather conservative. So, the result is an artificial extremism that fluctuates with the docket. But this is hardly the only problem with voting scores.

In addition to docket effects, the voting behavior of justices may reflect the political environment facing the Court. History suggests that the justices often moderate in the face of political opposition from the other branches (Rosenberg 1992). In my opinion, given a largely or completely endogenous docket and the possibility of strategic voting by the justices, it takes a lot of faith to see the justices' voting scores as a pure measure of their ideologies. In fairness, similar criticisms have been directed at the congressional DW-NOMINATE scores, but despite some rather intense work on the subject, no one has been able to show that docket effects or strategic behavior contaminate those scores in a profound way.

The bottom line is, for the nominees we need something other than their subsequent voting scores as a measure of their ideology, and we want this measure to be scaled like DW-NOMINATE scores. What to do?

Here is our solution. For the nominees who served in Congress, we start with their DW-NOMINATE scores. More specifically, five nominees (Black, Byrnes, Burton, Vinson, and Minton) served as legislators before they were nominated, so we can directly use their DW-NOMINATE scores as a measure of their ideology.<sup>6</sup> For those who did not, we start with the NOMINATE score of the president who nominated them.<sup>7</sup> This is hardly perfect, but it seems reasonable to believe that the ideology of the nominating president contains valuable information about the nominee. To complement this measure, we also employ the nominee ideology scores calculated by Jeffrey Segal and Albert Cover (1989), mentioned in the previous chapter. These scores are based on contemporaneous editorial evaluations in four major newspapers (two liberal, two conservative): *The New York Times*, the *Washington Post*, the *Chicago Tribune*, and the *Los Angeles Times*. Roughly speaking, they indicate the percentage of editorial evaluations that score the nominee as a liberal.

We believe that both sets of scores tap into the nominee's true ideology, though probably with error. To recover the posited common factor underlying the two measures, we performed a principal component analysis (PCA) on the two measures. Then, to put the PCA scores into the DW-NOMINATE space, we simply regressed them on the DW-NOMINATE measure associated with each nominee (their score or the president's score). Using the estimated coefficient, we then converted each PCA score into a DW-NOMINATE score. Essentially, we took a DW-NOMINATE score for each nominee (either his or her own or that of the nominating president) and then adjusted it based on contemporaneous perceptions of the nominee.

<sup>6</sup> Aficionados of ideology scores will note that we converted all of the scores into the Senate space.

<sup>7</sup> Because presidents announce their position on many bills, they have been scaled just like senators or House members.

Table 17.1. *Supreme Court nominees with scandals*

Nominees with hearings			
Scandal before hearing	Scandal during hearing	Scandal after hearing	No hearing
Haynsworth, Carswell, Rehnquist, 1 Bork, Kennedy, Breyer	Clark (last day), Fortas, 2 Rehnquist 2 (last day), Thomas	Jackson, Warren	Black

The recovered DW-NOMINATE scores have a great deal of face validity; that is, justices we know were liberals show up that way, and so do conservatives and moderates. These scores also do a pretty good job of predicting the future voting behavior of the nominees who make it onto the Court.

Finally, given the chair's score and the imputed score of the nominee, we calculate the simple distance between them (the absolute value). Given the range of observed DW-NOMINATE scores, the distances range between 0 and 1. The mean distance is .34 and the median is .31. The minimum distance is virtually 0 and the maximum is almost 1. The standard deviation is .2.

### Scandal

The theory indicates the importance of some dirt for the chair to work with. We will use the presence of a scandal as our measure of the stock of dirt. According to our data, coded from stories in *The New York Times*, thirteen nominees experienced some sort of scandal or scandals. Since we are interested in the effect of scandal on the length of hearing and the volume of the hearing report, the actual timing of the scandal is critical. We classified the nominees with scandal into three categories according to this timing; see Table 17.1.

We did not score late-breaking scandals – those that happened after the public hearings had concluded – since, for our purposes, we sought a measure of dirt existing at the time of the hearings. We did count the scandal involving Black, as it emerged between the time Roosevelt nominated him but before he was confirmed, which occurred without a hearing. A trickier case involves scandals that emerged during hearings. If one suspects that the chair engineered their emergence, or that they had an effect on the subsequent duration or intensity of the hearings, one should count them as dirt. But if one views them as nonengineered and not liable to affect the choreography of the hearings, one might not count them. So, one might count all four of the “hearing scandals” as scandals, or none of them, or just the two that occurred prior to the last day (indicated in Table 17.1). In our data analysis we opted for the first of these choices, but we reran all the results using the other coding choices, and it didn't make much difference to the results.

### Staff Size

For the size of the Senate Judiciary committee staff, we began with Keith W. Smith's compilation of data on congressional committee staff

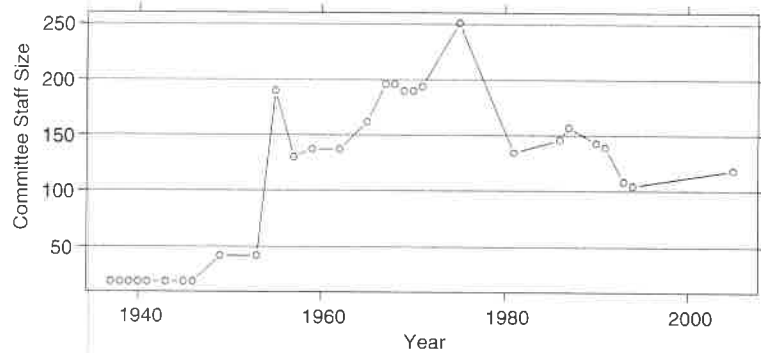


Figure 17.5. Size of the Senate Judiciary Committee staff at the time of Supreme Court nominations, 1937–2005.

(1947–98).<sup>8</sup> Smith compiled the data from standard sources. We have updated and checked these figures, using staff directories published by *Congressional Quarterly*. For pre-1947 nominations, for which we've had trouble locating exact numbers, we use the number for 1947. Figure 17.5 displays the data over time. As you can see, the size of the committee staff has varied greatly over time. Basically, staff increased steadily until the Stevens nomination in 1975 and then fell to a moderate and fairly stable level. So, this is not a story of simple growth.

### Senate Courtesy

Supreme Court nominations are usually referred to the Senate Judiciary Committee, which typically schedules a public hearing on the nominee. However, by unanimous consent, the Senate can directly consider and confirm a nomination without referral to the Judiciary Committee. This fast-track nomination process, called "Senate courtesy," has been invoked for a few nominees, invariably former or current senators. For example, Byrnes skipped the hearing due to Senate courtesy. Of course, Senate courtesy is hardly an inviolable rule (we find it hard to imagine that it would be invoked today). In fact, some former or current senators were nonetheless sent to the Judiciary Committee for hearings. However, in these cases, the nominee usually had a nominal hearing and the nomination was quickly sent to the full Senate for a vote. We identify Black, Byrnes, Burton, Vinson, and Minton as filling the historic requirements of Senate courtesy.

### VISUALIZING THE DATA

As we've emphasized repeatedly, visualizing your data is an essential part of any analysis and a cornerstone of modern applied statistics (Cleveland 1985, 1993). This is still true when you are testing a formal model, because you want to be fair

<sup>8</sup> Available on his personal Web site, <http://trc.ucdavis.edu/kwsmith>.

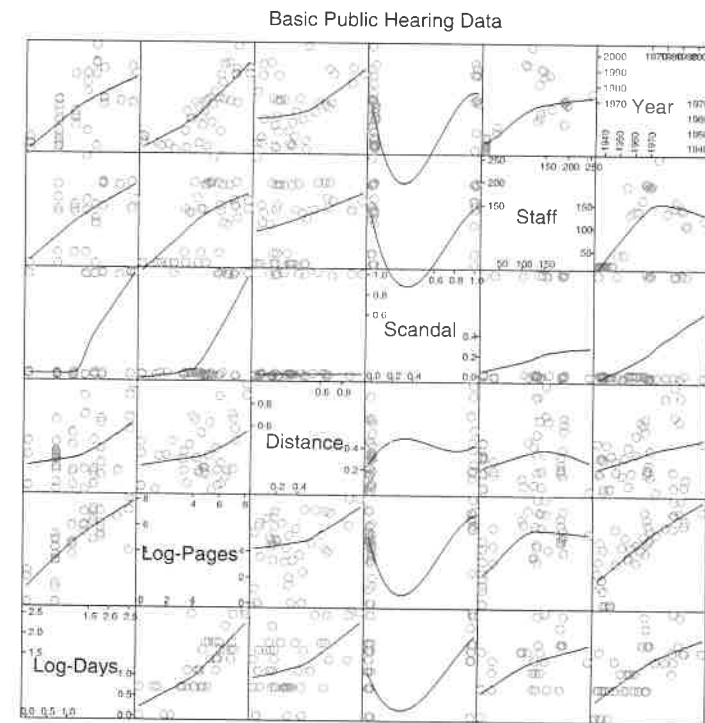


Figure 17.6. Scatterplot matrix of the data from the public hearings of Supreme Court nominees. Each dot represents a different nomination.

with the data. By this we mean that you want to be confident that the predicted patterns really exist. In addition, you want to know if the data contain interesting patterns that the theory doesn't predict. Surprises are great, because they are food for thought. They may even stimulate you to create a new theory. And there's nothing like clever visualizations for uncovering surprising patterns (and for finding coding errors or other mistakes, which happen all the time).

So, where do we start? I like a scatterplot matrix with nonparametric smoothers imposed on each panel: simple to do but often effective. This is shown in Figure 17.6.

Each row of Figure 17.6 has the indicated variable on the  $y$ -axis and each of the other variables on the  $x$ -axis. For example, look at the bottom row in the figure. The  $y$ -variable here is days of hearings (logged). Log-days is arrayed against each of the other variables in turn. It's nice to see log-days and log-pages track each other pretty closely, suggesting that both tap into the same thing, presumably hearing intensity. In fact, each of the key predictors (taken singly) appears to have a strong impact on the duration of the hearings and, in the next to bottom row, on log-pages. In the next to bottom row, you can see what looks like a threshold effect from staff on log-pages. Looking again at the bottom row, you may see something similar for duration, now that you look harder. In other words,



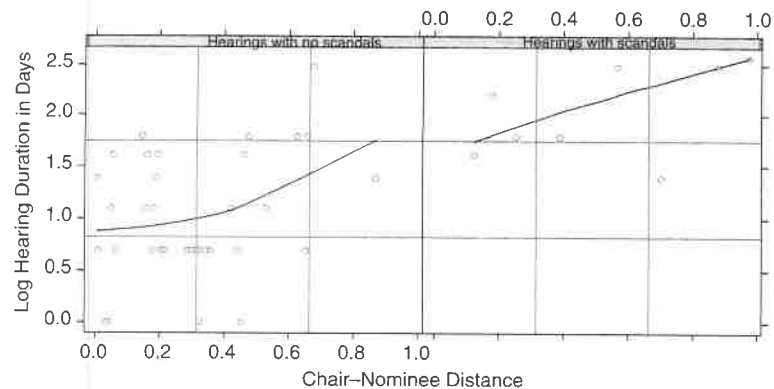


Figure 17.7. Hearing duration and ideological distance between the chair of the Senate Judiciary Committee and the Supreme Court nominee. The left and right halves of the graph show nominations without and with scandals.

once you get beyond a certain level of staff (130 or so), additional staff doesn't seem to have much effect. Now that we're a little sensitized to threshold effects, look again at the relation between distance and days and pages. This is interesting since the distance-duration relationship is central to the theory. Again, it looks like a threshold effect. This is gratifying, because our theoretical model predicted it (remember minimal hearings for proximate nominees?).

Examining the last column, you can see that most of the variables increase over time, though staff turns down again (we saw that earlier). This comovement over time is a little worrisome because it suggests that we may have trouble untangling all the relationships. I will return to this point later. Skimming the matrix for other strong relationships, I see that staff increases with the distance between the chair and the nominee. This suggests that the Judiciary Committee has more staff when it is estranged from the executive. It might be worth thinking harder about potential endogeneity, but we won't spend more time on this here.

So, despite a few worrisome issues, our initial cut at the data looks promising. What would be a potentially illuminating visualization, given our theory? Personally, I want to see the relationship between duration and ideological distance, for scandal-ridden nominees versus nonscandalous ones, and for staff-poor versus staff-rich nominations.

Figure 17.7 shows log-days as a function of the ideological distance between the chair and the nominee. The left panel shows the relationship for nominations in which no scandal emerged before or during the hearings. As shown, when the chair and the nominee are ideological soul mates (ideological distances under about .3 or .4 in DW-NOMINATE units), hearing durations are usually very short. Nor, in these circumstances, are durations particularly responsive to distance. But when the chair and the nominee are estranged (ideological distances greater than .4 or so), the durations get noticeably longer as ideological distance increases.

The right panel of Figure 17.7 shows data from nominations with scandals, and there the situation is even more dramatic. The hearing duration is much longer

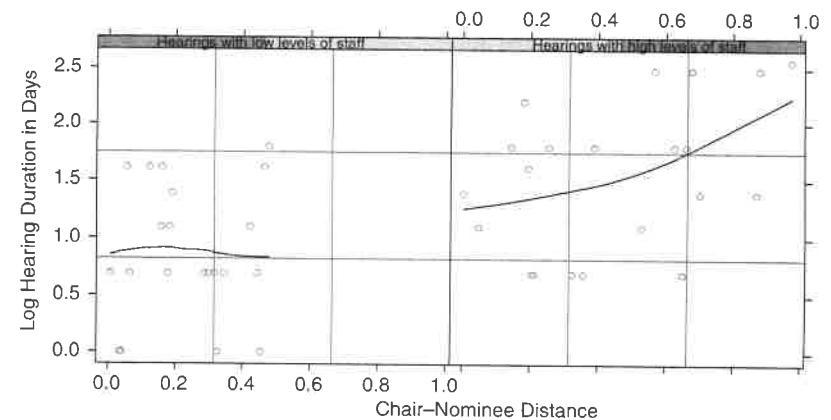


Figure 17.8. Hearing duration and ideological distance between the chair of the Senate Judiciary Committee and the Supreme Court nominee. The left and right halves of the graph show nominations with low and high levels of staffing.

even for ideological soul mates and goes up from there. In fact, the duration is longer for the closest scandal-ridden nominee than it is for the more distant scandal-free ones. Beyond the .4 mark, the slopes of the two lines don't really look that different, suggesting a merely additive effect of scandal rather than an interaction. One caveat worth noting: As the panels show, most of the scandal-free nominations involve ideologically proximate nominees. The scandal-ridden nominations span the ideological spectrum.

Now let's check out the effect of staff, as shown in Figure 17.8. Again, the figure shows the duration-distance relationship, but now the left-hand panel shows nominations in which the committee staffing was relatively skimpy (below 133, with a minimum value of 27). The right-hand panel shows the relationship for nominations in which staffing was relatively abundant (above 133, with a maximum of 251). The left half of the figure reveals that all the ideological distances are small. So, the figure says: When distances are small and staffing is tight, durations are typically short. It is tempting to add: "and they don't vary with distance." But we have to be careful here because of the limited range of the data. For really big distances the durations might tick up, but we can't tell since we have no observations in that range.<sup>9</sup> The right half of the figure has observations spanning the ideological spectrum. We see both an upward shift in the curve and a notably positive slope. Again, all this looks quite consistent with our theory.

<sup>9</sup> It is easy to be sloppy about projecting relationships outside the range of their data. Relying exclusively on tables of regression coefficients encourages this often risky practice. The defaults in the "trellis graphics" package I'm using won't project the smoothers outside the data, which is nice. More generally, when showing the fit from a regression, it is also a good idea to show the actual data points if possible or, barring that, a "rug" at the bottom of the figure with a dash for each data point, so that the reader can see the distribution of the data. Simple practices like this would contribute to more honesty (and self-awareness) in quantitative political analysis.

An obvious question now is: What happens if we vary staffing and scandal at the same time? If we had enough observations, we could use conditioning plots to examine both relationships simultaneously. Unfortunately, we don't have enough scandal-ridden nominations to do much with. However, we can take a closer look at the scandal-free nominations, and the patterns do not appear to change compared to the analysis of the full data.

### FITTING PARAMETRIC MODELS

We've learned a lot about the data just from plotting them in artful ways. So, let's go ahead and fit some parametric models. However, before we do so, we should ask ourselves: What class of models should we fit? There are at least three possibilities.

First, based on equation (3), we could fit log-days or log-pages using ordinary least squares. This is certainly the simplest option and, in light of what we saw earlier, we know it is likely to work well (there might still be heteroskedasticity with the pages data, though). However, there are several problems here. First, we have to decide what to do with the zero-duration nominations, since  $\log(0)$  is undefined. One choice is to use  $\log(\text{days} + 1)$ , but this is arbitrary and has no theoretical justification. Also, it is inconvenient to constantly have to convert back from log-days to days.

The second option is to treat the data as count data and model them appropriately. This means fitting the data with Poisson or perhaps negative binomial regressions. This approach automatically takes care of the zero-duration nominations, and we won't have to worry about converting back from log-days to days. Running count models is simple, so this is an attractive choice.

Third, some people view the hearing data as duration data. Accordingly, they would want to estimate survival curves and hazard rates. This is a perfectly sensible choice. But note that the theory does not say that the chair holds a hearing for a day, decides whether to stop or continue, depending on the realization of some random variable, holds it for another day if he or she continues, decides to stop or not, and so on. That would truly suggest a survival analysis. Rather, the theory says that the chair determines the length of the hearings *ex ante*, based on the information he or she has in hand. This suggests a count model<sup>10</sup> in which the expected values of the data are modeled on the log scale, with a probability distribution used to model the outcome, which can be 0, 1, 2, and so on.

To illustrate, let's cheat a bit and look ahead to Model 1 in Table 17.2, which shows the estimated coefficients (and standard errors) in a simple regression model. The mean values of "Ideological distance" and "Staff size" are 0.34 and 1.14, respectively. If we set those variables as their mean values and set scandal to

<sup>10</sup> The data are not quite durations in that the outcome is a count of "hearing days" on which the hearing was actually conducted. If we were to pursue the idea of a duration model, we would want a measure of the onset of each spell of hearings, noting each day a spell was in the risk set of termination. Be that as it may, we checked the statistical results of the Poisson and negative binomial regressions with a duration analysis, a Weibull parametric regression. The results of the Weibull analysis are essentially the same as those of the count models.

Table 17.2. Two simple models of hearing duration and volume

	Hearing duration (days)		Volume of documents (pages)	
	Model 1	Model 2	Model 3	Model 4
Ideological distance	1.1(.3)	1.1(.3)	1.6(.8)	1.7(.7)
Scandal	.6(.2)	.6(.2)	1.3(.5)	1.2(.4)
Staff size	.005(.002)	.005(.002)	.011(.004)	.007(.004)
Senate courtesy	—	-1.4(.8)	—	3.0(.7)
Constant	-.4(.3)	-.0(.3)	3.3(.5)	3.8(.5)
<i>N</i>	42	42	41	41
Pseudo <i>R</i> <sup>2</sup>	.18	.21	.04	.06
Log likelihood	-78.8	-76.1	-260.3	-254.0

*Note:* These models are overdispersed Poisson (i.e., negative-binomial) regressions with a logarithmic link so that coefficients can be interpreted on the log scale. For example, from Model 1, cases with and without scandals (with all other predictors unchanged) differ, on average, by .6 in their log duration. To put it another way, having a scandal corresponds to a duration that is, on average,  $\exp(.6) = 1.8$  times longer.

zero, we calculate  $E[y|\bar{\Delta}, \theta = 0, \bar{\kappa}] = e^{-.36+1.14(.34)+.61(0)+.01(1.14)} = 3.2$  days. This is the predicted duration of a scandal-free nomination, when "Staff size" and "Ideological distance" take their mean values. A similar calculation shows that the equivalent value for a scandal-wracked nomination is 5.9 days. You can see that the switch from no scandal to scandal increases the duration by  $(5.9 - 3.2)/3.2 = .84$  - in other words, an 84% increase. An approximation that is often used is that, roughly, the percentage change in  $E[y|x_1, \dots, x_n] = 100$  times the coefficient times the size of the change. So, for example, if ideological distance increases from 0 to 1, duration will go up 114%, that is, more than double (using the coefficients in Model 1). This approximation suggests about a 60% increase with the switch from no-scandal to scandal.

This all implicitly assumes a causal interpretation for the regression coefficients. To be more scrupulous, we would say something like "Comparing two nominations that differ in ideological distance but are identical in all other aspects, we would expect the nomination with ideological distance equal to 1 to have a duration 114% of that of the nomination of ideological distance 0." This is not the same as what would happen if the ideological distance of a particular case were to be changed (whatever that means). It is common to describe regression coefficients causally, but for observational data (which is almost always what we have in political science), all these coefficients actually tell us only about descriptive comparisons.

What would be a reasonable first-cut model, in light of the theory and our knowledge of the data's structure? The most straightforward answer is to model hearing length (days) and intensity (pages) as an additive function of ideological distance, scandal, and staff size. In addition, we should investigate senatorial courtesy. Finally, it would be a good idea to look at interactions among the important predictors (as discussed in the previous chapter).

Table 17.3. *Estimated coefficients and standard errors (all on the logarithmic scale) for overdispersed Poisson regressions of hearing durations (in days), with a distance threshold and a mid-1960s break*

	Model 2	Model 5	Model 6
Ideological distance	1.1(.3)	1.7(.4)	1.4(.4)
Scandal	.6(.2)	.6(.2)	.5(.2)
Staff size	.005(.002)	.005(.002)	.002(.002)
Senate courtesy	-1.4(.8)	-1.4(.7)	-1.3(.8)
Pre-1967	-	-	-.7(.3)
Constant	-.0(.3)	.2(.3)	.9(.4)
<i>N</i>	42	42	42
Residual deviance	43.0	41.6	35.1

Note: In Models 5 and 6 ideological distance is a threshold variable. Deviances can be compared to a null value of 115.

Take a look at the models in Table 17.2 of hearing length (days) and the volume of documents in the hearing. Model 1 is a plain vanilla model of the data. It contains just the three key variables: ideological distance, scandal, and staff size. Model 2 then adds senatorial courtesy. The three predictors are statistically significant at the 5% level (that is, the estimates are more than two standard errors from zero) and have the positive coefficient predicted by the formal model's comparative statics. Senate courtesy is not quite statistically significant. Models 3 and 4 show the corresponding results for the volume of the hearing report.

### ADDING A THRESHOLD

Our earlier discussion emphasized a threshold in distance as a distinctive theoretical prediction, one that received some support when we visualized the data. Let's fit a parametric model with a threshold to the days data and see how it compares to the earlier models. Here is a simple approach.<sup>11</sup> Let's say that we think that the threshold should be 0.4 on the distance scale. First, define an indicator variable that takes the value 0 if the distance is less than 0.4 and 1 if it is greater. Then interact this variable with (distance - 0.4). This procedure creates an elbow-shaped threshold variable: The variable takes the value of 0 at and below the threshold, then increases linearly (from 0) as distance increases above 0.4.

Let's estimate this model using the elbow version of distance and compare it with our simple model. Table 17.3 shows the results. The first column repeats Model 2 for purposes of comparison (that is, the distance measure is the simple absolute value used earlier). The second column shows the same model, but using the threshold version of distance. As you can see, the coefficients on the other variables have about the same magnitude (the intercept is different, of course).

<sup>11</sup> Another method is to estimate a so-called generalized additive model, with distance entered in a flexible way, say via smoothing splines (see Hastie and Tibshirani 1990). This gives similar results in this example.

However, the new distance measure is steeper (and more statistically significant). The overall fit of the model is modestly better. In sum, the data are equally compatible and perhaps favor the threshold version of ideological distance.

### WHAT COULD GO WRONG?

At this point, we could declare victory and go home. After all, we crafted a new theory of committee hearings, which we formalized in a mathematical model; we interrogated the formal model to uncover its empirical implications; we collected appropriate data to test the empirical predictions; we explored the data for structure using nonparametric models; we fit parametric models to the data; and both the parametric and nonparametric models behaved as the formal model predicted. Given this, we can now interpret the politics of public hearings on Supreme Court nominations in a way that is both theoretically informed and historically grounded. Admittedly, understanding the duration and intensity of public hearings on Supreme Court nominations is a small question, not like world peace or ending poverty. But still, in a modest way, we might call this a good day at the office.

But wait! We should ask one more question: What could go wrong with our pretty picture? Is there some way in which we haven't been completely fair with the data, something we overlooked perhaps because it didn't mesh with the theory?

I will leave the following as an exercise for you. Fit the model again, partitioning the data into early (pre-1967) and late (1967-on) groups. Since there are no senatorial courtesy nominees in the later period, you will have to drop this variable for the latter. You will find that the model works well in the later period (though staff is no longer significant). Even the coefficients look similar. But in the earlier period, the model seems to fall apart in that only the courtesy variable remains statistically significant.

What are we to make of this? There are two obvious interpretations. One interpretation is that the relationships elaborated in the theory hold throughout the period under study, but the effect of ideological distance is muted in the early period by the low values of distance in that period. Then, in the later period, we can detect it because distance takes on a wide range of values, all the way from 0 to 1. The second interpretation is that the world suddenly changed in 1967 (or thereabouts). Prior to that, the theory simply didn't apply because ideology was unimportant. Then suddenly, ideological conflict became critical in Supreme Court nominations. So, the first interpretation is that the evidence powerfully favors the theory after 1967, and we can understand why it would be hard to detect in the earlier period even if it held (the circumscribed range of the variables). So, we should assume that the theory holds throughout. The second interpretation is that the theory holds only in the last part of the dataset due to an (unexplained) structural break.

Let's see if we can get a little leverage on the two interpretations by examining residuals. Consider the top row of Figure 17.9. It displays the residuals from Model 5 in Table 17.3, our best model so far. The left panels of Figure 17.9 show the residuals for the early observations, arrayed against the (raw) distance measure. The right panel shows the residuals for the late observations. A lowess

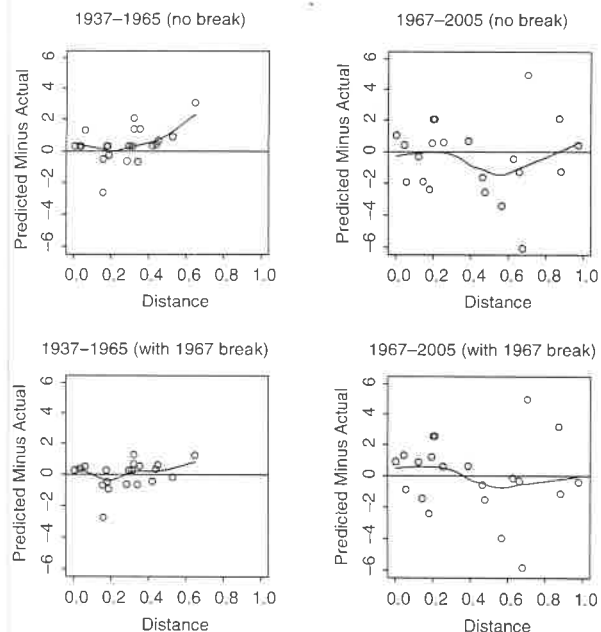


Figure 17.9. Early and late residuals for the Supreme Court model, plotted versus ideological distance, without (top row) and with (bottom row) a 1967 break

smoother shows the pattern in the data: The model tends to overestimate durations in the early period, especially if distance is greater than the threshold, 0.4. The model also tends to underestimate durations in the late nominations. This pattern indeed suggests a pre/post-1967 effect, but perhaps only a shift rather than a radical break.

So, let's fit that model. This is Model 6 in Table 17.3. There you can see that an indicator variable for whether a nomination came before 1967 is statistically significant. In this model, the staff variable is no longer statistically significant but this is not terribly surprising, given the time pattern in staffing. The bottom rows in Figure 17.9 display the residuals from this model. As you can see, the model fits the early observations much better. In the later period, the model no longer consistently underestimates the durations. However, it tends to slightly overestimate durations for nominations with small ideological distances and underestimate durations for those with large ideological distances. So, there remains some unexplained structure in the data, but Model 6 clearly outperforms the earlier ones.

The pre/post-1967 effect is completely ad hoc – it emerged from the data rather than from the theory. Of course, we can imagine some plausible reasons why 1967 might denote a watershed in the politics of Supreme Court nominations. We might point to the judicial activism of the Warren Court and the proliferation of new interest groups that occurred at that time. Controversial Supreme Court decisions coupled with burgeoning groups may have boosted the political salience of Supreme Court nominations, leading to greater publicity-seeking behavior on

the Judiciary Committee, beyond the simple effect of increased ideological distance.

So, it seems that we might want to tweak the theory to account for the varying political salience of the Court. In turn, this could lead us to collect systematic data on salience so that we can unpack the pre/post-1967 effect and see if our conjecture about salience and groups holds water.

I will stop here, but I hope you are getting a feel for the interplay between theory and data, a sort of back-and-forth tennis game between deduction and induction. Both halves are fun and worthwhile, but what I enjoy most is volleying back and forth.

### THE NORMATIVE ISSUES

In Chapter 15 I said that it was important to ask normative, evaluative questions once we had a solid foundation of theory and data upon which to stand. Well, we have a theory that works quite well, according to the data. What are the normative questions here? The model assumes that the chair of the Judiciary Committee tries to block ideologically distant nominees by exposing scandals. What are the implications of this model? Is this process a good thing?

Clearly, the process has some positive features. For one thing, turkeys get a good airing, at least when the nominee is ideologically far from the Judiciary Committee chair. So far as it goes, this is surely a good thing, as it lets senators know what they are voting for or against. It also makes it easier for citizens to evaluate their senators' votes and the president's performance as a chooser of nominees.

But I don't want to exaggerate. Even at the simplest level, the asymmetries in the process are disturbing. First, ideologically proximate turkeys tend to get a free pass in the Judiciary Committee, or at least a less costly passage. Hugo Black is an extreme case in point. So if you think that exposing turkeys is a good thing, then the manifest ideological bias in the process is a bad feature. If you feel this way, you might favor giving greater scope to the minority members of the Judiciary Committee, who will be inclined to be tougher on nominees the majority favors.

There is a second kind of asymmetry that many people find disturbing: the asymmetry between scandal and ideology. The model says that the real engine driving the process is ideology. But the hearings aren't forthrightly about ideology; instead, a hostile chair makes hay by publicizing the nominee's unsavory or shady behavior. Some people may see this behavior as trivializing or even misdirecting the process. The example of Black may be disturbing in this regard, as many observers believe he was an excellent justice despite his early membership in the Ku Klux Klan. But he surely would have had a tough time in hearings before a Republican-controlled Judiciary Committee and might well have been rejected by a Republican Senate. So, the emphasis on scandal may do a poor job of screening future justices.

The nasty way the hearings unfold during periods of ideological polarization may also contribute to public alienation from or disgust with government. Obviously, we haven't confirmed this conjecture with systematic data, and it

might be hard to do so. But the point does suggest itself. Still, even if this is true, it is probably an inevitable consequence of ideological polarization: Historically in America, polarized politics is nasty politics.

The process clearly creates incentives for the president, who can anticipate the action in the Judiciary Committee. First, it creates an incentive for the president to vet nominees and weed out the turkeys before they enter the hearing room, especially when the Judiciary Committee is hostile. Surely this is a good thing. It would be interesting if archival evidence showed this happening. Second, because of the way scrutiny is tied to ideology, the process may create an incentive for the president to avoid ideological extremists, at least when the chair is far from the president. By doing so, the president can take some of the steam out the hearings. If so, one might see this as a good feature of the process, especially in an age when most voters are moderates and most politicians are extremists. But we would need to investigate the president's choices in depth to see if this is really true before we start lauding the process for encouraging moderation.

Ultimately, we are concerned with a much larger normative question: Does the nomination process keep the Supreme Court within the ideological mainstream of American politics, and staffed with judges of intellectual competence and high integrity? For if the Court wanders outside the mainstream, it is apt to provoke potentially disastrous assaults on judicial independence. And a judiciary filled with corrupt, venal hacks will inevitably undermine the rule of law. The theoretical and empirical work in this chapter is only a very small step toward answering the big question we really care about. But the only way to build a cathedral (as it were) is by laying one brick on top of another, one at a time.

### Exercises

1. *Theory*. The formal model in Chapter 16 has implications about how many negative stories about the nominee we should see during the period of public hearings. To derive these implications, substitute equation (3) back into equation (2) and solve for the number of negative stories,  $n$ . Using this expression, derive comparative static predictions about  $n$ . Indicate what data you would need to test these predictions and what statistical methods you would use.
2. *Co-plots*. Form a conditioning plot similar to Figures 17.1–17.3 but conditioning on senatorial courtesy. Use the data shown in the scatterplot matrix to make two more conditioning plots of your choice.
3. *Senatorial courtesy*. Devise a logit or probit model that predicts when we should see senatorial courtesy. Think about the possible explanatory variables, such as ideological distance between the two parties. Do you believe your model?
4. *Analysis of the pages data*. Complete an analysis of the pages data similar to that in the chapter on the days data. In particular, fit negative-binomial and Quasi-Maximum Likelihood Estimation (QMLE) regressions after including the pre-1976 variable, and test whether a threshold version of distance is comparable with the data. (Refer to Table 17.1.)

## FURTHER READING FOR PART V

### 1. The Scope and Content of the Social Sciences

The ambitions of the social sciences. Historic development. Current disciplinary divisions. Warring tribes and allied coalitions. Enduring tensions.

#### REQUIRED READINGS

Ralf Dahrendorf, "Social Science," pp. 800–2 in *The Social Science Encyclopedia* (2nd ed.). This encyclopedia has a pronounced British and decidedly sociology-centric bent. Nonetheless, it is a fairly handy one-volume compilation of definitions and brief essays. This essay, by a distinguished sociologist, has some virtues.

#### OPTIONAL READINGS

Jon Elster, *Nuts and Bolts for the Social Sciences* (1989). If possible, read this and the new 2007 edition.

Dorothy Ross, *The Origins of American Social Science* (1991). The best intellectual history of pre–World War II American social science. But often it is tough going because of the author's postmodernist twitches.

Alain Desrosieres, *The Politics of Large Numbers: A History of Statistical Reasoning* (1998). A wonderful and social-scientifically sophisticated history of statistics and statistical concepts. The chapter on econometrics is particularly good.

### 2. How to Read Social Scientific Studies

Other courses focus on the tools you will need. But a quick and dirty introduction to some basics is necessary to understand the readings in this course. Part V covers theory–operationalization–data linkages. Basics of research design. Internal versus external validity. Small  $n$  versus large  $n$ : strengths and weaknesses. How to read and criticize regression results. The rhetoric of persuasion.

#### OPTIONAL READINGS

Donald T. Campbell and Julian C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (1963).

### 3. Case Study: Progress in the Social Sciences – How It Works

How, and to what extent, progress – not an unproblematic concept – is made in the social sciences. Role of evidence. Incentive structures for social scientists.

#### REQUIRED READINGS

Steven G. Brush, "Dynamics of Theory Change in the Social Sciences: Relative Deprivation and Collective Violence" (December 1996). The rise

and fall of a social scientific theory. Pay attention to the history of what happened and to the article itself as a piece of social science.

David Hull, "The Need for a Mechanism" (1988). This is a somewhat eccentric choice, but Hull offers extremely acute comments about how science works, including such phenomena as attribution and credit claiming, citation practices, the role of mechanisms and evidence in persuading others, insiders and outsiders, cooperation, competition, personal rivalries, and scientific fraud. Can you see what he is talking about in the history of "relative deprivation"? In addition, Hull sketches a social scientific theory of scientific progress.

#### 4. Overview of Political Science

The subject matter of political science. Subfields. Typical concerns. Important concepts. History of the discipline. The behavioral revolution. The rational choice revolution. Where the field is going.

##### REQUIRED READINGS

Robert Goodin and Hans-Dieter Klingemann, "Political Science: The Discipline" (1996). As your interest takes you, skim through the rest of the volume. The essays by Keohane, Barry, Weingast, Alt and Alesina, and Carmines and Huckfeldt are pretty good.

#### 5. Case Study: Pluralism and the "Economic" Theory of Regulation

Who wins and who loses from government programs? What is the government up to, anyway? Illustrates the rational choice approach in political science. Extremely practical ideas for understanding public policy.

##### REQUIRED READING

George J. Stigler, "The Theory of Economic Regulation" (1988).

##### OPTIONAL READING

David Truman, Introduction to *The Governmental Process: Political Interests and Public Opinion* (1993).

#### 6. Case Study: The Impact of TV News on Viewers' Political Behavior

How does TV news affect viewers' political beliefs and electoral behavior? Illustrates social psychological approaches, experimentation, and clever research design.

##### REQUIRED READING

Shanto Iyengar, "Experimental Designs for Political Communication Research: From Shopping Malls to the Internet" (1992). This article is

available online and includes some of the media manipulations used in Iyengar's recent work.

Shanto Iyengar and David R. Kinder, *News That Matters* (1987). The best book on television and politics.

#### 7. Case Study: The Activity Puzzle in Congressional Elections

Do the activities of congressmembers affect their reelection chances? It seems that they should, but the evidence says no; what's the problem here? Illustrates the kind of puzzle-solving that typifies the field of American politics. Also, the idea of selection bias: why it is pervasive and important in quasi-experimental data and one way to handle it.

##### REQUIRED READINGS

David Mayhew, *Congress: The Electoral Connection* (1973), a classic book on congressmembers' reelection activities.

Eric Schickler *Disjointed Pluralism: Institutional Innovation and the Development of the U.S. Congress* (2001).

Kenneth Shepsle and Barry Weingast, eds., *Positive Theories of Congressional Institutions* (1995).