Why Do Incumbent Senators Win? Evidence from a Dynamic Selection Model*

Gautam Gowrisankaran§ Matthew F. Mitchell¶ Andrea Moro‖

(PRELIMINARY) December 21, 2003

Abstract

We investigate why incumbents win disproportionately often. To do so, we structurally estimate the parameters of a dynamic model of voter behavior using U.S. Senate data. Our model specifies three potential reasons for the incumbency advantage: selection, tenure and challenger quality. Each of these separate effects is identified from data on histories of election outcomes. We estimate the parameters of the model using the method of maximum likelihood. We find that tenure effects are negative or small. The incumbency advantage is due to the average quality of incumbents being higher and to incumbents facing weaker challengers than candidates running for open seats.

*We thank Antonio Merlo, participants at the Society for Economic Dynamics 2003 conference and 2002-03 Industrial Organization graduate students at Harvard and Yale for their insightful comments.

§John M. Olin School of Business, Washington University in St. Louis and National Bureau of Economic Research. Email: gowrisankaran@olin.wustl.edu

¶Department of Economics, University of Iowa. Email: matthew-mitchell@uiowa.edu

‖Department of Economics, University of Minnesota. Email: amoro@atlas.socsci.umn.edu

1 Introduction

This paper investigates why incumbents win disproportionately often. In a variety of electoral situations, incumbents win substantially more than half of the time. This is sometimes referred to as an *incumbency advantage*; in the U.S. Senate, for instance, incumbents win approximately 75 percent of the time.

Many explanations have been posited to explain this fact, including pork-barrel spending, congressional relations, media coverage, incumbent visibility, and party attachment. One possibility is that some of the incumbency advantage is due to a selection effect: becoming an incumbent requires winning, and winning candidates will tend to be of relatively high quality. In this case the incumbency advantage is not due to any direct benefit of incumbency, but is rather a consequence of the different distribution of quality for incumbents.

The goal of this paper is to empirically disentangle the selection effect from other sources of incumbency advantage. Our method is structural. We formulate a simple model of voter decisions for candidates, estimate the parameters of the model, and use the estimated model to examine the relative importance of the selection effect to other sources of incumbency advantage.

In our model, voters obtain utility from their elected official through two sources. First, each candidate has a permanent, idiosyncratic quality, which voters value. Second, voters value the tenure of their elected official. There are several possible sources of tenure effects. Elected officials may learn on the job, they may use tenure to achieve better status within the Senate, or they may use the exposure of incumbency to their advantage in elections. It also could be that candidates become out of step with their constituents or their skills deteriorate over time. We do not attempt to separate the different sources of the tenure effects, as it is beyond the scope of our model or data. Instead, we simply seek to evaluate the importance and variation of tenure effects, and to separate them from the selection effect. We do not constrain the tenure effects to be monotonic in tenure.

We develop a stylized model of voters and candidates. Voters are identical dynamically optimizing agents, who do not differ in their preferences for policies or parties. Voters observe the permanent quality of two current candidates and then elect one of them. Permanent candidate quality is drawn from a fixed distribution, which varies depending on whether the election is an open seat election or an incumbent is running. Once quality is drawn, the only change in the utility flow from a candidate over his career is his tenure effect, which moves in an identical way across elected officials. An incumbent leaves the Senate with an exogenous exit probability that depends on tenure. As such, we do not account for the selection bias that may result from senators choosing when to retire based on their electoral prospects; we view this as reasonable in the vast majority of cases.

We estimate the parameters of our model using U.S. Senate data. The use of Senate data avoids several pitfalls present in other data sources. In particular, the U.S. House of Representatives contains many instances of redistricting and it is not clear how to treat elections following a redistricting, when two incumbents may run against each other.

Our dataset contains the history of senatorial seats, recording how candidates came to office, how long they served in office and the reason that they left office. Conditional on a given vector of fundamental parameters, the solution to the voter's dynamic choice problem implies a probability distribution over the possible electoral histories of a senatorial seat. We derive this distribution, and use it to estimate the parameters of the model with the method of maximum likelihood.

The model allows us to separately identify selection and tenure effects in a relatively intuitive way. The key feature of the data that allows identification is that two senators with identical tenure could have important differences in their electoral history and the winning probability of an incumbent depends on the entire history of the seat. For instance, a one-term incumbent could have beaten a two-term incumbent, or he could have beaten a five-term incumbent. If five-term incumbents almost never lose, then the econometrician's posterior distribution of permanent quality for the one-term incumbent who beats a five-term incumbent will be different from the posterior distribution for a one-term incumbent who beat a two-term incumbent. If selection based on candidate quality is an important determinant of the incumbency advantage, then these two incumbents would have quite different probabilities of reelection despite their identical tenure. Thus, the importance of selection will be identified based on how the electoral history affects reelection probabilities, conditioning on incumbent tenure.

To our knowledge, this paper is the first to estimate the role of selection in the incumbency advantage using an approach that is consistent with an optimizing model of voter choice and the first to consider the entire history of a seat in evaluating the role of selection. However, our approach is very parsimonious. For instance, we do not use the characteristics of candidates and elections, which have been used in other studies. We make this choice for two reasons. First, including additional characteristics would not be feasible within the context of a dynamic optimization model as these characteristics would be state variables. Second, we view the spirit of our study as providing evidence on the importance of selection compared to a mixture of all other effects, and not in providing a detailed account of the relative contribution of different factors to the advantage of incumbents.

Our results show that tenure effects are negative or small. Thus, there are two factors that explain the incumbency advantage. First, on average, incumbents are of higher quality due to selection effects. Second, on average, incumbents face a weaker pool of challengers than do candidates running for an open seat.

We performed several robustness checks of our model. Principally, we examined a variety of relevant moments for which the parameters are not explicitly chosen to fit. These moments appear to fit the data reasonably well at the estimated parameter values. For instance, the predicted incumbency advantage from the model is in line with the data. In addition, the winning probabilities as a function of tenure and the number of terms since the last open seat are close to the data. The model is capable of generating these high winning probabilities for incumbents without large tenure effects, demonstrating that selection effects are a big part of the story.

2 Relationship with the existing literature

Starting in the 1970s, a vast literature has tried to quantify incumbency advantages.[1] Early studies regressed the winning probability on an incumbency dummy. As we discussed above, interpreting a positive coefficient in this regression as a tenure effect is problematic because it may be due to a selection effect. The first method to try to separate the tenure and selection effects defined the tenure effect as the difference between the vote share that a senator earned in his second and first elections. This measure became known as the *sophomore surge*.[2] Gelman and King (1990) pointed out that the sophomore surge approach also suffers from selection bias because a candidate who is elected would disproportionately have had a good draw in his first election, that may be idiosyncratic to the first election.

[1] Most studies use House election data, which contains a larger number of elections. They typically regress winning probabilities on a set of regressors. See the references in the surveys by Cover and Mayhew (1977), Fiorina (1989), and Mayhew (1974). For more recent studies, see also Ansolabehere and Snyder (2002), Cox and Katz (1996), and Lee (2001), together with the other references cited in this section.

[2] See Erikson (1971), Cover (1977), Gelman and King (1990) and references therein.

They developed a reduced-form OLS method that helps mitigate this selection bias. Levitt and Wolfram (1997) apply a Heckman-style correction to the sophomore surge to further mitigate the Gelman and King (1990) selection bias. Separately, Levitt and Wolfram (1997) attempt to separate the sources of incumbency advantage into increased incumbent quality and decreased challenger quality, by considering pairs of elections where the same two candidates face each other.

Our paper builds on these earlier papers, in that we recognize, and attempt to control for, the problem of selection bias that is inherent in an electoral system. Our model incorporates the fact that an incumbent who won had a positive idiosyncratic shock in the first period, as in Gelman and King (1990). The idiosyncratic shock occurs because the winner of an open election likely faced a competitor whose quality was less than the population average, but will face an average competitor in his second election. As in Levitt and Wolfram (1997), we allow the new candidate density to differ based on whether the candidate is in an open seat election or not, in some specifications.

Our approach differs from the approach of these papers in several important ways. First, we estimate parameters that are fully consistent with an optimizing model of voter behavior. Thus, there are no longer any questions as to whether our estimates suffer from selection bias, conditional on our model being accurate. Second, our model is identified by the entire history of electoral outcomes since the open seat election, and not just by the data from the current election.[3] This allows us to identify our parameters of interest, without using the limited (and potentially biased) information available from when pairs of candidates face each other multiple times. Third, our model generalizes the sophomore surge approach, in that it allows for tenure effects, which we can estimate separately from the effect of selection, by using the entire history of outcomes. Fourth, we use data from the U.S. Senate, and not the U.S. House of Representatives. This allows us to avoid having to deal with the problems inherent in redistricting.[4] Fifth, do not include many of the covariates used in earlier studies. We made this choice because we need to keep the number of state variables to a minimum in order to estimate the dynamic model. Finally, we use data only on election wins, and not on vote shares. We made this decision in order to estimate parameters that are consistent with a well-specified model and because vote shares are more likely to be

[3]In this way, our model relates to Samuelson (1987), who first recognized the importance of the entire history of a seat in evaluating incumbency advantage.

[4]Redistricting is problematic because it is difficult to define incumbency and electoral outcomes in the case where one district is formed from parts of more than one district.

biased by endogenous challenger quality, which we do not model.

A related literature has structurally estimated candidate career decisions to retire or face reelection. Both strands of the literature attempt to predict reelection probabilities. Diermeier, Keane and Merlo (2002) estimate a model where candidate career decisions are endogenous but reelection probabilities are exogenous. In contrast, we endogenize reelection probabilities but treat retirement decisions as exogenous. We view our approach as complementary to a career decision model, as we seek to examine the extent to which endogenous voter decisions can explain reelection probabilities, but cannot also model endogenous retirement decisions.

3 The Model

We model the decision problem for a representative voter who values services from an elected official, in our case a senator. The valuation has two components: a senator-specific, permanent quality q and a tenure effect τ_m common to all senators of tenure m. The quality q is an element of a compact set Q. Tenure is defined by the number of completed terms in office. The utility flow for the voter in a given period is additive in these two components, i.e.

$$u(q, m) = q + \tau_m \tag{1}$$

The voter values the expected sum of current and future utility flows, discounted by β.

In each period, voters choose between two candidates in an election. There are two kinds of elections between which it is useful to distinguish. One is an *incumbent-challenger election*. This is an election where an incumbent runs against a challenger. The other type is called an *open seat election*, which takes place in situations where neither candidate is an incumbent. This happens when incumbents leave office for reasons other than losing an election. We assume that these reasons are exogenous and depend only on tenure.

The timing is as follows. At the beginning of the period, the incumbent either exits or runs for reelection. Denote the probability of exit at tenure m by δ_m. If he exits, two new candidates run for the seat. If he runs for reelection, a single challenger runs against the incumbent. Each new candidate then draws his permanent quality q from an atomless distribution $F(q)$[5] with corresponding density $f(q)$. In the estimation, we allow

[5] We assume that F is atomless to ensure that the voter has strict preferences over candidates with probability one.

F to be different depending on the type of election, to account for possible differences in the quality distribution of candidates running open seat elections vs. challengers running against incumbents. For simplicity of notation we abstract from this in the description of the model.

The incumbency effects τ_m are tenure-specific constants known to the voter. The voter observes the qualities of the current candidates and then elects the candidate that maximizes expected discounted utility. The voter also knows the distribution F from which future candidates will draw their permanent qualities.

For an open seat election, the optimal choice of the voter is simple: choose the candidate with the higher q. The utility flows generated by the candidates are otherwise identical.

In an incumbent-challenger election the decision is more complicated. We express the problem recursively using a Bellman equation. Denote by q the quality of the incumbent and by q_c the quality of the challenger. The state vector is the incumbent senator's quality q and tenure m. Let $V(q, m)$ denote the expected discounted utility for the voter at the beginning of the period, before either exit occurs or new candidates appear. Let W denote the expected discounted utility from an open seat. Then

$$V(q,m) = (1-\delta_m)\int_Q \max\left\{\begin{array}{c} q+\tau_m+\beta V(q,m+1), \\ q_c+\tau_0+\beta V(q_c,1) \end{array}\right\} f(q_c)dq_c + \delta_m W \qquad (2)$$

If the incumbent chooses to run again (with probability $1-\delta_m$), the voter chooses between the incumbent and a challenger. The integral in the first term in (2) reflects the expected utility in this case, which involves integrating over q_c.

If the incumbent exits, the voter obtains W. Letting the two challengers' qualities be defined by q and q_c,

$$W = \int_Q \int_Q \max\left\{\begin{array}{c} q+\tau_0+\beta V(q,1), \\ q_c+\tau_0+\beta V(q_c,1) \end{array}\right\} f(q)dqf(q_c)dq_c. \qquad (3)$$

The value of the open seat reflects the fact that two challengers are drawn and the higher q is retained.

Denote by $r(q, q_c, m)$ the optimal reelection rule of a voter when the incumbent has quality q and tenure m and the challenger has quality q_c; $r(q, q_c, m) = 1$ denotes reelection and $r(q, q_c, m) = 0$ denotes choosing the challenger. We now show that the solution to the decision problem can be characterized as a cutoff rule. As a result, the Bellman equation takes a simple form that is useful in computing the solution. We start by characterizing the decision rule.

Lemma 1 *$r(q, q_c, m)$ is weakly decreasing in q_c.*

The proof is in the appendix. The lemma implies that the voter follows a cutoff rule: challengers are elected only if their quality exceeds a cutoff $\bar{q}(q, m)$. Note that voters do not simply choose the candidate with the higher q, or even the higher $q + \tau_m$, since the voter is forward looking and considers future tenure effects and exit probabilities.

The cutoff rule allows us to express the Bellman equation more concisely. We do this by defining $V(q, 0)$ to be the voter's value function from an open seat election where the one candidate has drawn q and a second candidate's quality has yet to be drawn. Then, by letting $\delta_0 = 0$, equation (2) can be rewritten as

$$V(q,m) = (1-\delta_m)\max_{\bar{q}} \begin{pmatrix} F(\bar{q})\,(q + \tau_m + \beta V(q, m+1)) \\ + \int_{\bar{q}}^{\infty} (x + \tau_0 + \beta V(x,1))\, df(x)dx \end{pmatrix} \tag{4a}$$

$$+\delta_m \int_Q V(x,0) df(x)dx \tag{4b}$$

If the incumbent does not exit (the case given in 4a), the expected return has two components: first, the payoff when the incumbent is retained, times the probability of retention $F(\bar{q})$; second, the expected value of the challenger, conditional on his quality being above $\bar{q}$. If the incumbent exits, $\delta_m W$ can be rewritten as (4b) by using (4) to define $V(x, 0)$.

4 Estimation

4.1 Overview

Our goal is to provide inference on the fundamental parameters of our model: the candidate permanent quality density f, the incumbency effects τ_m, the exit probabilities δ_m, and the discount factor β. Our data contain information on when and how each U.S. senator came to office and when and how he left office. These data allow us to understand, for instance, whether a senator came to office by winning an open election or by defeating an incumbent.

We do not directly observe any component of quality. However, given a vector of fundamental parameters, the model generates a probability distribution over sequences of electoral outcomes. We use the method of maximum likelihood to find the parameter values that maximize the probability of seeing the observed electoral outcomes.

To understand how the model provides evidence on reelection probabilities that we observe in the data, it is useful to consider a special case. Suppose that incumbency effects and exit probabilities are constant across tenure, i.e. $\tau_m = \bar{\tau}$ and $\delta_m = \bar{\delta}$. In this case, the policy function satisfies $\bar{q}(q, m) = q$; the voter always chooses the candidate with the higher q because tenure does not affect current or future payoffs.

Suppose that candidate A won an open seat election in 1960 against candidate B and then defeated challenger C in 1966. After that election, we know that A's permanent quality q is distributed as the maximum of 3 i.i.d. draws from F. Suppose instead that C had won in 1966. Then, we can infer instead that C's permanent quality is distributed as the maximum of 3 i.i.d. draws from F. Thus, the probability of the incumbent winning in 1972 depends solely on the number of elections that have occurred since the last open election for that seat. As a result, the probability of reelection will be increasing in the number of terms since an open seat, and conditionally independent of tenure.

In general, the probability of reelection will depend on the entire history of wins and losses since an open seat election. Let us extend the electoral history of the previous paragraph to consider the 1972 election, and suppose that the challenger, D, wins in 1972. In the case where $\tau_1 = \tau_2 = \bar{\tau}$, our posterior on the permanent quality of D is independent of whether he beat A or C. However, consider the case where incumbency effects depend on tenure, for example $\tau_0 = \tau_1 < \tau_2$. For simplicity, assume that $\beta = 1$. If D defeated the two-term incumbent A in 1972, then D must have had a sufficiently high q to overcome his deficit in τ $(\tau_2 - \tau_1)$. In contrast, we cannot make the same inference if D defeated the one-term incumbent C in 1972. Thus, our posterior density of the permanent quality of D is higher if he beat A than if he beat C.

Extending this example to 1978, D's probability of being reelected depends not only on his tenure (1) and terms since an open seat election (3) but also on whether he beat A or C. This example demonstrates why the entire history matters.

The example also suggests how our model can separately identify incumbency effects from selection effects. Conditional on a candidate's tenure, the model will, for different parameter values, predict different probabilities of reelection given different histories since the last open seat election. By matching these predictions of the model to the data, we can understand the relative importance of selection and tenure effects.

This discussion also shows how difficult it would be to use a reduced-form approach to estimate tenure effects separately from selection effects. One cannot simply regress the probability of reelection on candidate tenure and simple statistics such as terms since an open seat or number of senators since an open seat. Any reduced-from regression would instead have to include the entire history since the open seat election, which grows exponentially in the number of elections since an open seat election. In our data set, this would imply thousands of regressors.

We now turn to the specifics of our data and our inference procedure.

4.2 Data and Institutional Background

We construct our dataset using data on U.S. Senate elections from the *Roster of U.S. Congressional Office Holders* (ICPSR 7803). In the original dataset each record refers to a senator seated in a given congress (a two-year period starting in odd-numbered years) and contains information about when and why the senator was seated, and when and why he left congress. The ICPSR dataset ends in 1998. We compiled more recent data in order to extend this dataset up to the 2002 election.[6]

We use these data to construct records of histories from an open election to an exit. We refer to one such history as a *chain*. Each chain is a vector of zeros and ones, with dimension equal to the number of elections held between the open seat election and the exit of the last senator in the chain. We do not include the outcomes of open seat elections in the chain. The first element of the vector is equal to one if the winner of the open seat election wins his next election. The second element is equal to one if the winner of the second election wins the third election, etc.

Normally the term of a senator lasts six years. Regular elections are held in November of even-numbered years. Seats are divided into three classes based on the year in which their regular elections are held. Senators take office in the January following their election.

Senators can leave office at the end of their terms essentially for three reasons, losing a general election, losing in the primary, and retirement. Our data contain instances where the senator leaves office before the end of his six-year term. Senators leave office before the end of their term because of death or retirement. In this case, an election is held on or before the next even-numbered November. The election is called a special election unless that senator seat was scheduled to have an election at that time. The governor of the state often appoints an individual to serve as senator until someone is elected.

[6]To gather the most recent data, we collected and compared information from various sources, including the Biographical Directory of the United States Congress (http://bioguide.congress.gov)

Every chain starts with an open seat. Open seat elections consist of all elections following the exit of a candidate because of death or retirement.[7] As a consequence, we treat all special elections as open seat elections even if one of the candidates briefly served as an unelected senator nominated by the governor. Our definition of an open seat election also implies that an election where the incumbent senator lost in the primary is not an open seat election. We treat the primary and election as a single election with two candidates.

We treat all elections, whether special or regular, as counting for one term. This simplification is imperfect because the time period in our model is one term. For instance, the voter discounts the future identically if there are four years between elections (due to a special election) or if there are six years between elections. Moreover, the interpretation of the tenure effects is that they depend on number of elections won rather than number of years served.

Senators have been elected by popular vote only after 1914.[8] Before this change, senators were appointed by the state legislature. As we do not have a model of how the state legislature chooses senators, we only consider data from elections held after 1914. Moreover, it is conceptually difficult to use chains that started on or before 1914, because we do not have a model for the density of permanent quality for an incumbent senator after 1914 unless every senator in his chain was elected and not appointed. Thus, our dataset contains only chains that start after 1914.

Our dataset contains 385 chains, with 576 different senators and 1330 elections.[9] We observe an exit preceding each of these 385 chains. Out of the 385 exits, 73 required a special election to choose the next senator. Considering all elections besides open seat elections, the incumbent senator won 735 out of 943 times (78 percent). Of the 201 incumbent losses, 43 occurred during the primary, with the rest occurring during the general election.

Some chains have dimension zero, which occurs when the winner of the open seat election exits without running for reelection; our dataset contains 82 chains of dimension zero. The chains contain at most 7 different senators and at most 15 elections. The longest tenure for a senator was Senator Strom Thurmond (R, SC), who served from 1954 to 2002, winning 8 elections. Only 8 senators served more than 6 terms. To avoid estimating parameters with very few observations, we assume that $\tau_m = \tau_6$ and $\delta_m = \delta_6$ for all $m \geq 6$.

[7] We observe cases where a senator loses an election and then retires between the election and the end of his term. We ignore the retirement decision in these cases.

[8] U.S. Constitutional Amendment XVII.

[9] 19 senators served non-contiguous terms.

4.3 Inference and Likelihood

As is well known in the literature, it is difficult to estimate the discount factor of a dynamic discrete choice problem (see Rust, 1987, and Magnac and Thesmar, 2002). Given that a regular term lasts six years, we set $\beta = 0.96^6$. In principle, we could jointly estimate all of the other parameters. However, since we treat the retirement probability as exogenous, we can obtain consistent estimates of the retirement probabilities δ_m without solving the voter's decision problem. Specifically, we estimate δ_m as the number of senators that retire with tenure m divided by the total number of senators that held office for at least m terms.

We allow the quality distribution of new candidates to vary depending on the type of elections. Specifically, we assume that candidates in an open seat election draw permanent quality q from a distribution F_o, whereas challengers of incumbents draw their permanent quality from a distribution F_c. We assume that both distributions are normal with means μ_o and μ_c and variances σ_o and σ_c, respectively. Note that μ_o and μ_c are not separately identified given our data: a shift in both means would not change any observable prediction of the model. Thus, we normalize $\mu_c = 0$ and estimate μ_o. Similarly, we cannot separately identify σ_o from σ_c, but can identify σ_o/σ_c. Thus, we normalize $\sigma_c = 1$. Although σ_o is then identified from the data, we also normalize $\sigma_o = 1$ in the interest of parsimony.[10]

One of the tenure effect parameters, without loss of generality τ_o, is not identified: adding a constant to τ_0 and the same constant to $\tau_1, ...\tau_6$ would yield the same predictions. Therefore, we set $\tau_0 = 0$.

We now discuss the estimation of the other parameters. Consider first the contribution to the likelihood of chain d of dimension T. Denote the history of wins and losses prior to the t^{th} election in the chain with the vector $h_t \equiv \langle d_1, ...d_{t-1} \rangle$.[11] Denote the posterior density over incumbent quality after history h_t as $g(\cdot|h_t)$, and the number of terms served by the incumbent holding office after history h_t as m_{h_t}. Define e_t as the random variable that indicates the outcome of the tth election in the chain, with the interpretation that $e_t = 1$ indicates the incumbent winning the election and $e_t = 0$ indicates the incumbent

[10] We estimate a specification where we constrain $f_o = f_c$. For this specification, we set $\mu = 0$ and $\sigma = 1$ as these parameters are not identified.

[11] Recall from our definition of a chain the first element, d_1, is the outcome of the re-election of the senator winning the open seat election. Hence this vector does not contain the outcome of the open seat election, which have no informaional content from our purposes.

losing. We can then express the likelihood as:

$$
\begin{aligned}
L(d|\tau_1,...,\tau_6) &= \prod_{t=1}^{T} \Pr\left(e_t = d_t|h_t\right) && (5)\\
&= \prod_{t=1}^{T} \int \left\{d_t \cdot F_c\left(\bar{q}(x, m_{h_t})\right) + (1-d_t)\cdot\left[1 - F_c\left(\bar{q}(x, m_{h_t})\right)\right]\right\} dg(x|h_t)dx.
\end{aligned}
$$

The expression (5) depends on the policy function $\bar{q}$, which in turn depends on the parameters. The expression also depends on the density of permanent quality for the incumbent at the start of period t, g_{h_t}. We evaluate this density using Bayes Law and the policy function $\bar{q}$. Call p the "prior" density of the incumbent at time t and decompose history h_t into two elements: the outcome of last period election d_{t-1} and the previous history h_{t-1}. Bayes Law implies

$$
\underset{\text{Posterior density given } d_{t-1}}{g(q|d_{t-1};h_{t-1})} = \frac{\underset{\text{Prior density}}{p(q|h_{t-1})} \cdot \underset{\text{Probability of outcome } d_{t-1} \text{ given } q}{\Pr\left(d_{t-1}|q;h_{t-1}\right)}}{\underset{\text{Unconditional probability of outcome } d_{t-1}}{\Pr\left(d_{t-1}|h_{t-1}\right)}} \tag{6}
$$

The prior density p is equal to f in the case where the incumbent was a challenger in the previous period, and is defined recursively as equal to $g(\cdot|h_{t-1})$ otherwise.

We now show how this formula is applied to three specific cases. First, consider the density of a one-term incumbent who won an open seat election in the previous period. In this case the prior density is f_o, and the conditional probability of winning the open seat given q is $F_o\left(\bar{q}(q,0)\right)$. For this case, (6) can be written as:

$$
g\left(q|h_0\right) = \frac{f_o\left(q\right)\cdot F_o\left(\bar{q}(q,0)\right)}{\int_Q f_o\left(x\right)\cdot F_o\left(\bar{q}(x,0)\right)dx}. \tag{7}
$$

Next, consider the cases with $t > 1$. We distinguish two cases, depending on whether the incumbent won or lost in the previous election. If $d_{t-1} = 1$, then the conditional probability of the election outcome in the previous period is equal to the probability that the challenger draws a permanent quality less than the threshold value $\bar{q}(q, m_{h_{t-1}})$, hence:

$$
g\left(q|h_t\right) = \frac{g\left(q|h_{t-1}\right)\cdot F_c\left(\bar{q}(q,m_{h_{t-1}})\right)}{\int_Q g\left(x|_{h_{t-1}}\right)\cdot F_c\left(\bar{q}(x,m_{h_t-1})\right)dx}, \tag{8}
$$

Finally, if $d_{t-1} = 0$, the incumbent was a challenger at $t-1$. This means that the prior density is f and that the his permanent quality q is greater than the threshold $\bar{q}(\cdot)$ implied by the voters' decision rule, which is a function of the previous incumbent's quality and

δ_1	0.1507
δ_2	0.2269
δ_3	0.2850
δ_4	0.3300
δ_5	0.3043
δ_6	0.5652

Table 1: Exit probabilities by tenure

history. Since the previous incumbent quality is distributed according to $g(\cdot|h_{t-1})$, then equation (6) can be written as:

$$g\left(q|h_t\right) = \frac{f_c\left(q\right) \cdot \int\limits_{z:\bar{q}(z,m_{h_{t-1}})<q} g(z|h_{t-1})dz}{\int_Q f_c\left(x\right) \cdot \int\limits_{z:\bar{q}(z,m_{h_{t-1}})<x} g(z|h_{t-1})dzdx}, \tag{9}$$

In order to evaluate the log likelihood of our dataset for a given parameter vector, we first compute the policy function using numerical dynamic programming. We then evaluate the likelihood for a chain using the computed policy function, together with (5) and (6), and sum the log of the likelihood for each chain. Details on the numerical procedure used in the estimation are in the appendix.

5 Results[12]

Table 1 shows the exit probabilities (computed directly from the data) that we used in the estimation. Table 2 shows the estimation results. Model 1 refers to the case where all candidates are drawn from the same distribution; in Model 2, we allow for the possibility that open seat candidates draw from a distribution with a different mean from candidates challenging incumbents.

The bootstrapped standard errors are in parentheses. Recall that the relevant units are that the standard deviation on the distribution of quality is one. The results suggest that any advantage to incumbency is not inherent to the office, but rather the result of the weaker candidates running as challengers against incumbents. The inherent disadvantage of incumbency measured for incumbents in Model 2 is small; for the most common cases,

[12]The results reported in this section are still preliminary.

	Model 1 ($f_o = f_c$)	Model 2 ($f_o \neq f_c$)
τ_1	0.005 (0.14)	-0.618 (0.10)
τ_2	0.055 (0.18)	-0.586 (0.12)
τ_3	0.125 (0.28)	-0.654 (0.25)
τ_4	-1.050 (0.59)	-1.290 (0.39)
τ_5	1.209 (1.58)	0.028 (0.86)
τ_6	-0.336 (0.40)	-1.126 (0.36)
$\mu_o - \mu_c$	n/a	0.724 (0.05)
logLik	-512.557	-491.445

Table 2: Estimated parameters (standard errors in parenthesis)

incumbents of one, two, or three years of tenure, the incumbency disadvantage τ is estimated to be less than one half of one standard deviation of the quality of challengers. The point estimate suggests that candidates running in an open seat election are superior to challengers who run against incumbents; as a way to see the magnitude, consider that the average candidate in an open seat election would be in the 75th percentile of the quality distribution for challengers to an incumbent.

Our point estimates for Model 2 reveal that tenure makes it harder to remain in office, hence the incumbency advantage is due to a combination of selection and incumbents facing weaker challengers. We ask how big the incumbent's benefit from facing weaker challengers are. We obtain statistics from the model by simulating election outcomes, and compare results computed using the policy evaluated at the maximum likelihood parameter values of Model 2 to results computed by assuming $\mu_o = \mu_c$. In the data, incumbents win 78 percent of the time. In our estimated model, they win 82 percent of the time. However, if incumbents were faced with challengers of the same quality as in an open seat election, then they would win only 64 percent of the time.

We also use the simulated data to evaluate the performance of the model. A simple statistic by which to judge the model is average terms served. In our data, the average terms served by a senator is 1.73. In the base model, it is 1.67, whereas in the Model 2 it is 1.74.

Another useful measure of goodness of fit is to see how well the model reproduces reelection probabilities. Since we are fundamentally interested in differences in reelection probabilities as a function of tenure, consider the following comparison of reelection proba-

Tenure	Model 1 ($f_o = f_c$)	Data	Model 2 ($f_o \neq f_c$)
1	0.72	0.75	0.76
2	0.79	0.80	0.81
3	0.80	0.81	0.84
4	0.76	0.79	0.82
5 or more	0.88	0.92	0.94
Average	0.77	0.78	0.80

Table 3: Goodness of fit: reelection probabilities by tenure

bilities, by tenure, in the model and data, displayed in Table 3.

Note that the model with identical distributions for all challengers, Model 1, underestimates the winning probability of one-term incumbents. One way to interpret this is that there is something special about one-term incumbents that is hard for the model to replicate. Model 2 adds exactly that feature, taking account of this through the higher mean value of the distribution from which challengers for open seats draw. It would be troubling to argue that incumbency effects are small in a model that underpredicts reelection probabilities, and hence does not generate sufficient incumbency advantage in winning probabilities. Model 2, however, overpredicts winning probabilities with small (or even negative) incumbency effects. In other words, the model is not throwing out incumbency effects at the expense of generating a high winning percentage for incumbents; it generates them all through the selection effect.

We suggested earlier that other variables, such as the number of terms since the last open seat, are relevant to the reelection probabilities because of the information that these variables convey about the selection effect. Conventional "sophomore surge" analysis either looks only at the case where the number of terms since the last open seat election is one, or lumps together all one term incumbents. Our approach attempts, and some cases is successful, in taking advantage of information contained in the variation across electoral situations. Table 4 illustrates how well our estimated parameters fit the conditional probabilities of reelection we observe in the data jointly in terms since last open seat election and tenure.

Once again we see clear evidence for different distributions for challengers to open seats and incumbent-challenger elections. In the case where the incumbent has just won an

Terms of Tenure \ Terms since last open seat	1	2	3	4	5
1	Model 1: 0.67 Model 2: 0.77 Data: 0.79	0.74 0.70 0.69	0.80 0.74 0.65	0.82 0.74 0.59	0.87 0.77 0.65
2		0.75 0.82 0.77	0.80 0.76 0.75	0.84 0.82 0.86	0.87 0.82 0.91
3			0.77 0.85 0.80	0.80 0.82 1.00	0.85 0.82 0.82*
4				0.72 0.85 0.81	0.79 0.75 0.83*
5					0.93 0.94 0.95

Table 4: Goodness of fit: reelection probabilities by tenure and terms since open seat, (* indicates fewer than 10 observations))

open seat, model one underpredicts the winning probability for the incumbent, but model 2 explains it well. One feature that our model cannot explain is the descending values generated by the data for column one.

6 Discussion and Conclusion

We analyze the causes of incumbency advantage for the U.S. Senate by structurally estimating a dynamic optimizing model of voter behavior. Our results are identified by examining the impact of the entire history of election outcomes following an open election on the probability that an incumbent wins, conditioning on the tenure of the incumbent. We find that the incumbency advantage is due to two effects. Incumbent senators are, on average, selected to be of high quality, due to their past successes in winning elections. In addition, incumbent senators are able to deter high quality challengers. However, we find no evidence of other benefits intrinsic to incumbency: tenure appears to provide a small disadvantage in reelection.

Our result that tenure is not an important determinant of the incumbency stands in contrast to some of the literature. However, studies of incumbency advantage have mostly focused on the U.S. House of Representatives. Relative to the House, the importance of tenure in the Senate may be less, because committee assignments are not as important.

Our result that incumbents face weaker challengers does have precedent in the literature. For instance, Levitt and Wolfram (1997) find the same effect in House elections, although with a different identification mechanism. Note that the selection effect still implies that incumbent senators are hard to beat, and therefore our finding that potential challengers might be dissuaded from running against incumbents is not puzzling.

There is substantial information available about elections besides the history of a seat that might be relevant, but that we do not use. For example, we could allow the density of permanent quality for a candidate to vary based on his previous career path. In particular, House experience could affect a candidate's quality by adding to his legislative experience. However, among senators who had just won an open seat election, those with House experience were reelected 79.8 percent of the time, while the figure was 78 percent for those without House experience. We did not use House experience because it does not appear to be an important predictor of the reelection probability.

In addition, the vast majority of the literature on incumbency effects uses the party of the winning candidates as a control variable. We investigated whether party membership has an effect on the reelection probability. In open seat elections, candidates from the same party as the previous senator win only 39 percent of the time. Similarly, among senators who have just won an open election, incumbents from the same party as the previous incumbent win 71 percent of the time, versus 83 percent of the time for different party incumbents. Thus, if anything, there appears to be a small negative effect of past party affiliation. It is conceptually difficult to model a negative correlation in the preference for parties across chains. Thus, we did not use party affiliation.

In spite of the fact that we do not include information on career experience or party affiliation, our estimated model generates conditional reelection probabilities that are close to those in the data. The main objective of this study is not in accounting for the relative contribution of different sources of tenure effects, but rather, in describing the extent to which selection is an important determinant of incumbency advantage. Selection appears to be an important determinant. We hope that future research can further quantify the magnitudes of different sources of selection and tenure effects.

A Appendix

A.1 Proof of Lemma 1

Proof. Denote by M the finite set of allowable tenures for an incumbent. We prove the lemma by first showing that $V(q, m)$ is increasing in the first argument, using standard recursive techniques (c.f. Stokey, et al. (1989)). Denote by C the metric space of all continuous functions $g : Q \times M \to \mathbb{R}$ that are weakly increasing in the first argument, where the metric is defined by the sup norm. Note that C is a complete metric space. Define the mapping T for any function $g \in C$ by

$$\begin{aligned} T(g) &= (1-\delta_m) \int \max \left\{ \begin{array}{c} (q + \tau_m + \beta g(q, m+1)), \\ (q_c + \tau_0 + \beta g(q_c, 1)) \end{array} \right\} df(q_c) dq_c \\ &\quad + \delta_m \int \int \max \left\{ \begin{array}{c} (q + \tau_0 + \beta g(q, 1)), \\ (q_c + \tau_0 + \beta g(q_c, 1)) \end{array} \right\} df(q) dq df(q_c) dq_c \end{aligned}$$

Notice that, whenever g is weakly increasing and continuous, so is $T(g)$, so T is an operator, $T : C \to C$. Notice that T meets Blackwell's sufficient conditions for a contraction: for any function $g' \geq g$, $T(g') \geq T(g)$, and for a constant c, $T(g+c) = T(g) + \beta c$, $0 < \beta < 1$. Hence, by the contraction mapping lemma, for all functions $V_0 \in C$, the sequence defined by $V_n = G(V_{n-1})$ converges to a function $V \in C$ that is the unique fixed point of the operator T. Since V is the fixed point, it is exactly the value function that solves the dynamic programming problem. Since $V \in C$, the value function V is weakly increasing in the first argument.

Note that q_c only shows up in two places in the choice of candidates (once in the current reward to choosing the challenger, once in the future value if the challenger is chosen), and if V is weakly increasing in the first argument, the total discounted reward from choosing the challenger is strictly increasing in q_c, while the discounted reward from choosing the incumbent is constant in q_c. As a result, $r(q, q_c, m)$ must be decreasing in q_c for any fixed q and m. ■

A.2 Details on the numerical computation

The permanent quality distribution F is continuous, which implies that we need to approximate the value function in our computation. We choose a discrete grid approximation, and use 301 evenly spaced grid points between -6 and $+6$, in order to capture the tails of the standard normal density. We use linear interpolation in order to create a smooth policy function $\bar{q}$, necessary for an efficient search for the maximum likelihood parameters.

We find the parameter vector that maximizes the likelihood by using numerical search algorithms. We use two different algorithms: a routine that we developed that combines the simplex method with random jumps and the method of simulated annealing by Goffe, Ferrier and Rogers, (1992).

We compute standard errors and confidence intervals using a bootstrap method. To perform the bootstrap, we resample from the data with replacement, in order to create a dataset with the same number of chains as our original dataset. We then evaluate the maximum likelihood parameter estimates with the new dataset. We repeat this procedure 20 times in order to find a density of the parameter estimates, which we use to derive the standard errors and confidence intervals.

References

[1] Ansolabehere, Stephen, and James M. Snyder Jr., "The Incumbency Advantage in U.S. Elections: An Analysis of State and Federal Offices, 1942-2000", Unpublished working paper, MIT Dept. of Political Science, 2002.

[2] Cover, Albert D., "One Good Term Deserves Another: the Advantage of Incumbency in Congressional Elections", *American Journal of Political Science,* 21, 3, 523-41, August 1977.

[3] Cover, Albert D. and David R. Mayhew, "Congressional Dynamics and the Decline of Competitive Congressional Elections", in *Congress Reconsidered,* eds Lawrence C. Dodd and Bruce I. Oppenheimer, New York: Praeger, 1977.

[4] Cox, Gary W., and Jonathan N. Katz, "Why did the Incumbency Advantage in the U.S. House Grow?", *American Journal of Political Science,* 40, 2, 478-97, May 1996.

[5] Diermeier, Daniel, Michael Keane and Antonio Merlo, "A Political Economy Model of Congressional Careers", PIER Working Paper 02-029, July 2002.

[6] Erikson, Robert S., "The Advantage of Incumbency in Congressional Elections", *Polity,* 3, 395-405, 1971.

[7] Fiorina, Morris P. *Congress: Keystone of the Washington Establishment.* Second Edition. New Haven: Yale University Press, 1989.

[8] Goffe, William L., Ferrier, Gary D and Rogers, John, "Simulated Annealing: An Initial Application in Econometrics," Computer Science in Economics & Management, Vol. 5 (2) pp. 133-46, 1992.

[9] Gelman, Andrew and Gary King, "Estimating Incumbency Advantage without Bias", *American Journal of Political Science,* 34, 4, 1142-64, November 1990.

[10] Lee, David S., "The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the U.S.", NBER Working paper #8441, August 2001.

[11] Levitt, Steven D., and Catherine D. Wolfram, "Decomposing the Sources of Incumbency Advantage in the U.S. House," *Legislative Studies Quarterly* XXII, 1, 45-60, February 1997.

[12] Magnac, Thierry and David Thesmar "Identifying Dynamic Discrete Decision Processes" *Econometrica,* 70, 2, 801-16, March 2002

[13] Mayhew, David R., *Congress: The Electoral Connection,* Yale University Press, 1974.

[14] Rust, John "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, 55, 5, 999-1033, Sep. 1987.

[15] Samuelson, Larry, "A Test of the Revealed-preference Phenomenon in Congressional Elections", *Public Choice* 54, 141-69, 1987.

[16] Stokey, Nancy, and Robert Lucas with Edward Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.