.

Estimating Incumbency Advantage: Evidence from Three Natural Experiments*

Yusaku Horiuchi

Crawford School of Economics and Government

Australian National University, ACT 0200, Australia

yusaku.horiuchi@anu.edu.au

http://horiuchi.org

Andrew Leigh

Economics Program

Research School of Social Sciences

Australian National University, ACT 0200, Australia

andrew.leigh@anu.edu.au

http://andrewleigh.org

Late Updated: October 2, 2009

* Prepared for presentation at the University of New South Wales on October 14, 2009. We thank xxx for research assistance and yyy for financial assistance.

Abstract

Do incumbent legislators have a relative advantage in getting votes as compared to challengers? In estimating this so-called "incumbency advantage," researchers have attempted to use various natural experiment techniques. Yet such papers typically use only one approach and, accordingly, estimate one local average treatment effect (LATE) at a time, making it difficult to interpret the average treatment effect (ATE) in a given electoral setting. To address this, we present results from three natural experiments in the same elections; specifically, focusing on vote-share discontinuity, boundary discontinuity, and random ballot ordering in Australian Lower House elections. We discuss the strengths and weaknesses of each approach, and how the LATEs might compare to the ATE for all incumbents. The results suggest that the estimates of incumbency advantage are sensitive to identification strategy. It also suggests, however, that incumbency advantage in Australia is smaller than previous estimated using data from the United States.

Key word: incumbency advantage, natural experiment, local average treatment effect, Australia

JEL classification:

1. Introduction

Do incumbent legislators have a relative advantage in getting votes as compared to challengers? Estimating the size of this so-called "incumbent advantage" has been a staple of empirical political science for several decades. In this paper, we re-visit this frequently-studied subject using new data from Australia. Specifically, by comparing the results from three different natural experiments in the same elections, we discuss the strengths and weaknesses of each approach and how the three local average treatment effects (LATEs) might compare to the average treatment effect (ATE) for all incumbents.

Incumbents may gain an advantage from several sources (see e.g., Ansolabehere et al 2000). First, incumbent members may provide services to their district (e.g., regulatory advice, federal grants), for which they are rewarded at the ballot box. Second, incumbent members may receive more votes simply because of their name recognition. Third, the typical incumbent may face lower quality challengers, with the most promising candidates from the opposing party biding their time until the seat is open.

Researchers have investigated the sources and magnitude of such an electoral advantage not only out of their theoretical concerns but also for normative reasons. With a very large incumbency advantage, legislatures will have many experienced hands, but little fresh talent. Conversely, no incumbency advantage (or an incumbency penalty) might raise the opposite risk: plenty of fresh talent, but too few experienced politicians. Given the normative ground, studies of incumbency advantage have often been connected to policy debates on term limits, campaign regulations, and entry requirements for new challengers.

Although the concept of incumbency advantage may sound intuitive and straightforward, what still remains controversial is how to estimate it properly. Simple regression techniques are said to produce biased estimates, because many attributes, which are correlated with why a particular legislator was elected (i.e., became an incumbent) and how well he/she performs in the next election, are difficult to measure or unobservable. These omitted or immeasurable variables include candidates' strength to raise campaign resources, the density of social networks they use when mobilizing votes, the party's policy attractiveness, to name a few. Incumbency advantage should be estimated after controlling for attributes specific to candidates and parties, as well as a range of other factors. This is, however, not an easy task.

To address this methodological issue, the existing literature has used a variety of natural experiment techniques. For example, Levitt and Wolfram (1997) focus on repeat contests, in which the same pair of individuals contest an election once when neither is the incumbent, and then again when one is the incumbent. Another approach is that of Ansolabehere et al (2000), who focus on redistricting, in which the same set of voters find themselves shifted from one electorate to another. And a third strategy implemented by Lee (2008) is to use regression discontinuity in the vote share, effectively comparing incumbents who won the previous race with 50+ε percent of the vote with challengers who lost the previous race with 50−ε percent of the vote.

A crucial point about each of these techniques is that they estimate different local average treatment effects (LATEs) – the average effects of a treatment variable for specific sub-populations. This naturally raises questions about the generalizability of these findings. For example, to the extent that election victory is quasi-random (as will be true when we move closer to the 50 percent breakpoint) and thus unrelated to other confounding factors, the regression discontinuity design will validity estimate the causal impact of incumbency on vote share at the next election. However, this LATE may not generalize to incumbency effects for elections that are *not* close. If candidates who lose by small margins become embittered, while those who win by small margins are grateful to the electorate, the local average treatment effect of incumbency may be larger for narrow winners than for the typical incumbent.

One way of coping with this problem is to compare across multiple natural experiments, using the same set of elections. This has the advantage that we are able to generate multiple LATEs and to delve more deeply into the strengths and weaknesses of particular empirical approaches for estimating incumbency advantage. We aim to extend the existing literature both by comparing multiple approaches for estimating incumbency advantage, and through a more explicit focus on the particular parameter that we are estimating.

In particular, our analysis uses three natural experiments to estimate the advantage that accrues to an incumbent party; namely incumbent *party* advantage.[1], First, we implement the vote share regression discontinuity approach of Lee (2008), effectively comparing the vote share of candidates who were narrow winners and narrow losers at the previous poll. The approach is becoming increasingly common in the economics and political science literature (see xxx 200x for a review), and we add to this literature using previously under-investigated data from Australia. Australia provides an interesting case for comparison with the U.S. (Lee 2008), because both countries have similar institutional settings – single-member districts and effectively two-party competition.

[1] Our methodologies are not well-suited to separating party and candidate incumbency advantages. A question we thus intend to examine is: "From the party's perspective, what is the electoral gain to being the incumbent party in a district, relative to not being the incumbent party?" (Lee 2008, 692).

Second, we use a geographic regression discontinuity technique, restricting the sample to shared polling booths that sit close to the boundary of two adjacent districts. The intuition behind this approach is that we are able to compare two sets of voters who live in close proximity to one another – one set has an incumbent from a given party, while the other does not. Geographical discontinuity methods have been used in other contexts (e.g., Black 1999; Middleton and Green 2008; see Davidoff and Leigh 2008 for a review).

The third natural experiment uses random variation in ballot ordering, which is an important feature in Australian Lower House elections. This allows us to compare the vote share of candidates who received a better ballot draw in the previous election with candidates who received a worse ballot draw. Holding constant the number of candidates, ballot order is truly a random variable, so it represents the cleanest of our three natural experiments. There are some studies estimating ballot order effects (Ho and Imai 2006, 2008; King and Leigh 2009), but we are not aware of any study using the ballot order as an instrument.

The remainder of our paper is structured as follows. In section 2, we outline our Australian data and institutional context. In section 3, we present incumbency advantage estimates from close elections. In section 4, we show estimates using shared polling places. In section 5, we present estimates using random ballot order. The final section discusses the results and concludes. The results of three experiments suggest that the estimates of incumbency advantage are sensitive to identification strategy. It also suggests, however, that incumbency advantage in Australia is smaller than previous estimated using data from the United States. This is consistent to some recent comparative studies on incumbency advantage, which may suggest that the U.S. is exceptional.

2. Data and Institutional Context

Australia is a bicameral parliamentary democracy with single-member districts in the House of Representatives (Lower House) and multi-member districts with state/territory boundaries in the Senate (Upper House). Voting is compulsory, and ballots are counted using preferential voting (also known as instant runoff voting in the House of Representatives and Single Transferable Vote in the Senate). Our focus in this paper is the Lower House. Re-election rates in Australian Lower House elections are slightly lower than those in U.S. Lower House elections. For example, 92 percent of Australian incumbents seeking re-election in the 2004 election were successful (130/141), while the incumbent re-election rate was 99 percent (399/401) in the 2004 U.S. elections.

At the national level, there are effectively two political parties: the left-wing Australian Labor Party (ALP), and a right-wing Coalition of the predominantly urban Liberal Party of Australia (LP) and the rural National Party of Australia (NP). Election dates are chosen by the government, with a maximum term of three years. During the period of our analysis, federal government was held by the ALP (which governed from 1983 to 1996) and the Coalition (which governed from 1996 to 2007). The 2007 election saw the ALP win office.

Our dependent variable is the two-party preferred (TPP) vote share, which is the share of votes after preferences have been distributed to two major parties (ALP and NP/LP) from independent and minor party candidates.[2] To avoid the problem that the two-party vote share of

[2] In our second experiment focusing on shared polling places, we use the two-candidate preferred (TCP) vote share because the TPP vote share is not available at the level of polling place. The TCP vote share is the share of votes after preferences have been distributed to the top two

the two major candidates in the same district must sum to 100 percent, our dataset contains only the vote share of ALP candidates. Naturally, it makes no substantive difference to our results if we use Coalition candidates instead.

For the purpose of estimating the incumbency advantage, Australia has three important advantages. First, electoral districts ("electoral divisions" in official documents in Australia) are drawn by a non-partisan committee (relevant for our second approach). In the case of the U.S., partisan redistricting gives room to gerrymandering. Methodologically, this implies that a natural experiment design is political contaminated. This is not the case in Australia. Second, as we already explained, ballot ordering is random and thus we are able to have a truly random instrumental variable (for our third experiment). Finally, voting is compulsory. We can minimize a contaminating factor that voters non-randomly choose not to go to the polls (and thus not to vote for an incumbent party) depending on the expected effects of an incumbent party repeatedly elected.

3. Using close elections to estimate incumbency advantage

In estimating the causal impact of incumbency on vote share, the fundamental problem is that we would ideally like to compare candidate i 's performance as an incumbent, y_{1i} , with her performance as a challenger, y_{0i}.[3] If we could observe both y_{1i} and y_{0i}, then the (individual)

---

candidates from others. Since the top two are almost always an ALP candidate and a Coalition (either LP or NP), the TCP vote share is equal to the TPP vote share in nearly all instances.

[3] As we noted, our methods allow us to estimate the incumbent *party* advantage. Thus, a winner (i.e., an incumbent legislator) or a loser in a given electoral district in the previous election and a candidate from the same party in the same district in the current election may not necessarily be

incumbency effect, A_i, would simply be the difference between them: $A_i = y_{1i} - y_{0i}$. The problem is that we never observe both y_{1i} and y_{0i} for the same candidate. Formally, we observe $y_i = y_{0i} + z_i(y_{1i} - y_{0i})$ where z_i is an indicator variable of the incumbency status. Therefore, in any empirical test of a causal hypothesis, we can only estimate the average effect; most typically, the average treatment effect (ATE) across all legislators:[4]

$$ATE = \frac{1}{N}\sum_i (y_{1i} - y_{0i}) \tag{1}$$

Since it is impossible, however, to assume that *all* legislators are determined by the toss of a coin or a draw from a hat, we need to find a "natural experimental" situation, in which the treatment (incumbency status) is assigned at (nearly) random at least for a subset of legislators so that we can estimate the "local" average treatment effect (LATE). Our three strategies address this fundamental problem effectively.

The first empirical strategy is to use a regression discontinuity design, focusing on the vote share of candidates who narrowly win or lose in the previous election. For this experiment, we use district-level data from 1984 to 2007. Specifically, the dependent variable Y_{it} is the ALP's vote share in district i in election t, and the treatment variable Z_{it-1} is whether or not a the ALP won a seat in district i in previous election $t-1$.

---

the same. In most cases, the same individual is the incumbent, but there is more turnover of individuals among challengers.

[4] Other average effects include the average treatment effect for the treated (i.e., incumbents) and the average treatment effect for the untreated (i.e., challengers).

Following the notation of Imbens and Lemieux (2008), we refer to the vote share in the previous election, X_{it-1}, as a "forcing" variable, since it determines whether or not an ALP candidate is the incumbent (i.e., whether or not the ALP won a seat). If $X_{it-1} > 0.5$, the candidate is the incumbent in the current election. If $X_{it-1} < 0.5$, the candidate is the challenger in the current election.

The intuition of regression discontinuity is that as one comes closer to the 50 percent breakpoint, the winner is increasingly likely to be determined by luck than by other factors such as skill or resources. In the limit, omitting subscripts for convenience, the incumbency advantage (expressed $A^{(1)}$ as LATE for our first experiment) can be as follows:

$$A^{(1)} = \lim_{x \downarrow 0.5} \mathrm{E}[Y \mid X = x] - \lim_{x \uparrow 0.5} \mathrm{E}[Y \mid X = x] \tag{2}$$

In theory, if it were possible to compare candidates precisely at the 50 percent breakpoint, then it would be the case that we could simply estimate the incumbency advantage by comparing the vote share of incumbents, $Y_{it}(1) \equiv Y_{it}(Z_{it} = 1)$ and challengers $Y_{it}(0) \equiv Y_{it}(Z_{it} = 0)$:

$$A^{(1)} = \mathrm{E}\{[Y(1) \mid X = 0.5] - [Y(0) \mid X = 0.5]\} \tag{3}$$

In practice, however, we observe no such elections in our data. We must therefore make additional assumptions that for both challengers and incumbents, the vote share in this election is a continuous function of the vote share in the previous election.[5] Under these assumptions (or

[5] Specifically, these assumptions are the following: For challengers, $E[Y(0) \mid X = x]$ is continuous in x and $F_{Y(0)|X}(y \mid x)$ is continuous in x for all y. For incumbents, $E[Y(1) \mid X = x]$ is continuous in x and $F_{Y(1)|X}(y \mid x)$ is continuous in x for all y.

indeed under the weaker assumption of continuity at $X_{it-1} = 0.5$, the incumbency advantage estimated in equation (3) is equivalent to the incumbency advantage estimated in equation (2).[6] As Imbens and Lemieux (2008) put it, "the estimand is the difference of two regression functions at a point" (p. xx).

Empirically, we proceed first by graphing the data, fitting local linear regressions separately for challengers (i.e., those with a vote share in the previous election of less than 50 percent) and for incumbents (i.e., those with a vote share in the previous election of more than 50 percent). Specifically, we run a linear regression using the vote shares in the current election Y_{it} and the previous election X_{it-1} for many subsets of observations. The total number of observations is 1,411. We then connect the predicted values, draw a line, and see whether we observe any discontinuity at $X_{it-1} = 0.5$.[7]

Figure 1 shows the results of this analysis.

[Figure 1 about here]

Not surprisingly, the two variables are strongly correlated. The higher the vote share is in the previous election, the higher the vote share in the current election. Consequently, the two predicted lines are almost along the 45-degree line. These patterns are, however, unimportant for our analysis. Our focus is the difference in the predicted values evaluated at $X_{it-1} = 0.5$. The

6 This is because $E[Y(0) \mid X = 0.5] = \lim_{x \uparrow 0.5} \mathrm{E}[Y(0) \mid X = x] = \lim_{x \uparrow 0.5} \mathrm{E}[Y \mid X = x]$ and $E[Y(1) \mid X = 0.5] = \lim_{x \downarrow 0.5} \mathrm{E}[Y(1) \mid X = x] = \lim_{x \downarrow 0.5} \mathrm{E}[Y \mid X = x]$.

7 For local liner smoothing, we use the rectangle kernel function.

figure suggests that there is indeed a positive incumbency effect, but the magnitude of the effect seems to be small.

To estimate the confidence interval of the gap at the discontinuity and to test the robustness of our findings, we did the following additional analysis. First, we choose three different bandwidths for local linear regressions – the width of the smoothing window around each point. The larger bandwidth implies the larger number of observations included in each local regression. This improves efficiency at the risk of breading down the balance between two groups – i.e., breaking the as-if random assumption near the discontinuity. For sensitively analysis, therefore, we need to estimate the LATE with different bandwidths. Specifically, we use the default bandwidth estimated by STATA 11's **lpoly** command, its 50%, and its 150%. Second, for each bandwidth, we estimate the size of the gap in Y_{it} at $X_{it-1} = 0.5$. Finally, we estimate the (normal-approximation) confidence interval based on bootstrapping with 50 replications.

Table 1 shows the results.

[Table 1 about here]

The estimated effects of incumbency advantage ranges from 1.007 to 1.735. The standard errors are large and, thus, the all estimated effects are not significant at the conventional level. Interestingly, compared with Lee (2008), these are substantially smaller impacts: In the U.S., the effect is at around 8 percent and statistically significant (Figure 4, p. 688). Why do the two democracies with the similar nature of electoral competition (i.e., two-party competition in single-member districts) have different incumbent advantage remains a puzzle. A possible explanation is that resources, particularly campaign funds and staff members, incumbents can use are much larger in the U.S. than in Australia. Another explanation may related to the fact that voting is compulsory in Australia but not in the U.S. American voters who support an incumbent in close competition are encouraged to go to the polls, while those who support a challenger abstain even when race is highly competitive. Some information gaps between two groups of supports may explain the difference in voter turnout near the discontinuity.[8] These hypotheses are worth examining further in more in-depth comparative analysis.

4. Using shared polling places to estimate incumbency advantage

Our second empirical strategy focuses on shared polling places and estimates the magnitude of the incumbency advantage at around *geographical* discontinuity. Specifically, using polling-place-level data from 1998 to 2007, we look at the differences in the ALP's vote share in bordering polling booths that serve two electorates.[9]

Formally, one can think of the forcing variable X_{jit} now being the distance from the polling booth j for district i in election t to the nearest electoral boundary. On one side of the boundary, $X_{it} < 0$, and when approaching the boundary, $X_{it} \uparrow 0$. On the other side of the boundary, $X_{it} > 0$, and when approaching the boundary, $X_{it} \downarrow 0$. Suppose that a polling booth sits precisely on the boundary (which is not uncommon, since booths are often located in schools, while boundary lines frequently run down major arterial roads). In this case, the incumbency

---

[8] In fact, in the U.S. case, the level of voter turnout may not be similar near the discontinuity and it may cause biased causal estimates. It is worth replicating Lee's study and testing the robustness of his findings by adding the turnout variable.

[9] The detailed information about polling stations (e.g., address) are available only for 1998 elections onward.

advantage (expressed $A^{(2)}$ as LATE for our second experiment) can be expressed by a similar equation to equation (3):

$$A^{(2)} = \mathrm{E}\{[Y(1) \mid X = 0] - [Y(0) \mid X = 0]\} \tag{4}$$

In this formulation, equation (4), denoting a bordering booth, is an analogous case to the one in which an election is decided by the toss of a coin (though of course bordering booths are a less "perfect" case than this).

The intuition behind such an approach is that if a main road marks the boundary, then it is likely that voters on each side of the road would have had similar voting patterns, but for the fact that those on opposite sides of the road have a different incumbent politician. Since it seems unlikely that individuals choose which side of the road to live based on the electoral boundary, any observed differences in voting behavior likely reflect the impact of incumbency on voting patterns. Another important feature in the Australian context is that unlike the U.S., an important feature of electoral politics in Australia is that electoral boundaries are drawn by a nonpartisan body, the Australian Electoral Commission. Therefore, we can assume that electoral boundaries are drawn irrespective of who is an incumbent on each size of the road.

Intuitively, this strategy has some similarities with Ansolabehere et al (2000), who exploit redistricting as a means of estimating the causal impact of incumbency. However, our approach differs in that we do not directly exploit changes in boundaries. Instead, as with the vote share discontinuity approach, our analysis is based on the assumption that voters living in close proximity to one another (so close that they cast their ballots in the same polling booth) would have voted in the same manner, but for differential incumbency effects.

Our regression specification for this second strategy is the following. The dependent variable Y_{ijt} is the ALP's vote share in polling booth j for district i in election t. The treatment variable Z_{ijt} is 1 if the ALP had a seat as of election t and 0 otherwise. Obviously, the incumbency status is the same for all j within a given i, but it can be different between districts for a given shared polling place j. The model also includes other covariates (W_{ijt}) and polling-place-year fixed effects (u_{jt}) and, thus, a full functional form is specified as follows:

$$Y_{ijt} = \beta \cdot Z_{ijt} + \delta \cdot W_{ijt-1} + u_{jt} + \varepsilon_{ijt} \tag{5}$$

where β approximates to $A^{(2)}$ as long as the incumbency status is well balanced between observations *within* a shared polling place in a given year, conditional on observable covariates W_{ijt-1}. In order to avoid potential bias, for W_{ijt-1}, we add a set of dummy variables for the number of candidates in district j in election $t-1$ and a set of dummy variables measuring which party was the major opponent for the ALP in district j in election $t-1$; namely, the LP, the NP, the Australian Greens (GR) or an independent.[10] These pre-treatment variables may explain whether or not an ALP candidate won in the previous election, but their correlation with the ALP's vote share in the current election is expected to be weak given the fixed effects.

It is important to note that we can powerful control a range of demographic covariates with polling-station fixed effects and that such an analysis can be done only when we have a sufficiently large number of shared polling places. It is equally important to note, however, that since we focus on variations within shared polling places (in specific years), our results are only identified from instances in which the same polling booth serves multiple electorates. In other words, we estimate the LATE of the incumbency status for candidates who compete for votes within a small area near the electoral boundary.

[10] Three candidates and independents are base categories, respectively.

We estimate four models. Models 1 and 2 include all 30,092 observations for 1998-2007 elections. The number of panels (polling-place/year) is 28,710. Therefore, 95 percent of observations are data from polling places *not* shared by multiple districts. Since there are four elections covered in our data (1988, 2001, 2004, and 2007), the average number of observations from shared polling places in each year is 345. Most of these observations are polling places shared by two districts. Models 3 and 4 include 1,449 observations (729 panels), where the share of total votes within each polling place is larger than 10% or smaller than 90%. Therefore, these models include non-shared polling places, as well as polling places shared but predominantly for a single electorate. Although the number of observations for estimation is dramatically reduced, this selection is expected to balance between treated observations (i.e., with an ALP incumbent) and untreated observations (i.e., without an ALP incumbent) within shared polling places. Models 1 and 3 are based on un-weighted OLS regressions, whereas Models 2 and 4 are weighted by the total number of votes (which is about 90% of the total number of eligible voters in Australia) in polling places. Since the denominator of the dependent variable ranges from 2 to 7,145,[11] it is preferable to run weighted least-square regressions to cope with the problem of heteroskedasticity.[12]

The results are presented in Table 2.

[Table 2 about here]

The magnitude of incumbency advantage is 9.025 (un-weighted) or 6.350 (weighted) with all observations. By restricting observations to shared polling places, it becomes 5.683 (un-

[11] These are among all observations (for Models 1 and 2). The mean is 1,270.

[12] We also use clustered robust standard errors where clusters are polling places-years. We thus assume that observations are independent across panels but may be correlated within panels.

weighted) or 5.348 (weighted). All of them are statistically highly significant. Note that additional covariates – the number of candidate dummies and the opponent dummies – tend to be significant in Models 1 and 2, but not significant in Models 3 and 4. This implies that they are not substantial determinants of the ALP's vote share within shared polling places. Given that restricting observations tend to improve balance (in other words, dropping causally irrelevant observations by a method equivalent to matching), we are inclined to conclude that our second LATE is about 5-6 percent. This is larger than the first LATE focusing on close elections but is still smaller than the estimates using the U.S. data.

5. Using ballot order to estimate incumbency advantage

Since 1984, Australia has used a random draw to assign ballot order. In line with research from the U.S. and U.K. (e.g., Ho and Imai 2006, 2008, xxx 200x), King and Leigh (2009) find that for a major party candidate, drawing the top position on the ballot yields an increase in the vote share of approximately 1 percentage point.

Our third estimation strategy focuses on this random variation and calculates the magnitude of the incumbency advantage by estimating an IV regression, in which the ballot order in the previous election is used as an instrument for incumbency status. This is perhaps the most ideal natural experimental setup, as we have a truly random variable. As long as the ballot order has a sufficiently strong correlation with the incumbency status, we can validity estimate another LATE for incumbents who luckily won in the previous election with an advantageous ballot position. Conceptually, the incumbency advantage in this analysis (expressed $A^{(2)}$ as LATE for our third experiment) is:

$$A^{(3)} = \mathrm{E}\{[Y(1) \mid X = x] - [Y(0) \mid X \neq x]\} \tag{6}$$

where an instrumental variable X_{it} for an ALP candidates in district x in election t is whether it is a certain favourable ballot position (x) or not.

A problem is that we do not know, a priori, which position is the most advantages for candidates to win a seat. The previous studies suggest that it is the first one, but these findings do not necessarily preclude the possibility that some other positions (say, the second and third ones) are also "good" positions, particularly when the number of candidates is large, which is the case in Australian Lower House elections. Therefore, we include a set of dummies for all ballot positions for ALP candidates in the previous elections. Considering the possibility that their opponent's ballot positions may also matter for their winning probability, we also include a set of dummies for ballot positions for Coalition (LP or NP) candidates in the previous elections.

The regression model for the third experiment is specified as follows:

$$Y_{it} = \beta \cdot \hat{Z}_{it} + \delta \cdot W_{it-1} + u_t + \varepsilon_{it} \tag{7}$$

where the dependent variable is the ALP's vote share in in district x in election t (=1987, …, 2007). $\hat{Z}_{it}$ is the predicted incumbency status based on the first-stage regression, which is estimated with the two sets of ballot-order dummies mentioned above and W_{it-1} and u_t. The former is a set of dummies for the number of candidates in the previous elections and the latter is election-specific fixed effect. Since the probability of having a particular ballot position is obviously conditional on the total number of candidates, we also include a set of dummies for the number of candidates in the previous elections. Since the number of candidates in the previous election may also correlate with the outcome variable, we treat W_{it-1} as included instruments.[13]

[13] We do not include dummies for the number of candidates in the current election because they are causally posterior to our treatment variable.

Following Chamberlain and Imbens (2009), we estimate this IV specification using both standard two-stage least squares (2SLS) and limited-information maximum likelihood (LIML). In a simulation using randomly assigned quarter-of-birth dummies, the authors show that LIML performs substantially better than 2SLS, particularly when instruments are not strongly correlated with the treatment variable. For comparision, we also run a standard OLS regression.

The results of OLS, 2SLS and LIML regression analysis are shown in Table 3.

[Table 3 about here]

The estimated incumbency advantage is almost the same in all the three models – 17.823 (OLS), 17.753 (2SLS) and 17.631 (LIML). All these effects are highly significant. We should be cautious, however, in interpreting these results, because the ballot order may suffer from the well-known "weak instrument" problem, in which there is a weak correlation between the excluded instrument and the endogenous regressor. The partial R-squared statistic of excluded instruments in the first-stage regression is only 0.021 and the F-statistic for the joint significance of these instruments is only 1.13. Considering a possibility that we add too many (weak) instruments, we also attempted a range of possible combinations (without any theoretical ground) of excluded instruments. No combination, however, yield a sufficiently large (typically, larger than 10) value for the first-stage F-statistic. Given these weak instruments, it is unsurprising to see instrumental-variable estimates being biased toward the OLS estimate, which is roughly the difference between the ALP's vote share for incumbents and the ALP's vote share for challengers. As Figure 1 suggests, this difference is large, but it may not imply that there is large incumbency advantage.

It is intuitively straightforward to see why our instruments are extremely weak. Although King and Leigh (2009) estimate that around 7 percent of contests were decided by a margin that was smaller than our estimated effect of being placed first on the ballot (1 percentage point), it does not follow that ballot ordering changed the result of 7 percent of races. If ballot ordering operates primarily through a first-position effect, it will typically be the case that neither major party candidate draws top spot on the ballot. King and Leigh, therefore, estimate that the first-position effect changed the result in only 1 percent of races. While it is plausible that this is a lower bound (our analysis allows for the possibility that ballot order makes a difference for lower-ranked candidates), it is plausible that our instrument only has "bite" for around 1 in 100 candidates.

6. Discussion and conclusion

What can we learn from comparing across methodologies? First, the magnitude of incumbency advantage is sensitive to the approach used. To the extent that the true incumbency effect is a convex combination of these approaches, researchers are more likely to come to a correct answer if they employ multiple approaches.

Second, if we discard the results from the ballot order experiment (which we are inclined to do), the remaining estimates of the incumbent party effect are almost nil (from the vote share discontinuity approach) and about 6 percent (from the bordering booths approach with restricted samples). We are inclined to think that the true Australian incumbency effect lies between these estimates, suggesting that the incumbency advantage is smaller in Australia than in the United States. This would be consistent with past research from other countries, which has typically found smaller incumbency effects than for the United States or even negative effects (see e.g., Gaines 1998; Linden 2004; Uppal 2009).

References

Ansolabehere, Stephen, James M. Snyder, Jr., and Charles Stewart, III. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Journal of Political Science*, 44(1): 17-34.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 114(2): 577–599.

Booth, Alison L. and HiauJoo Kee. 2009. "Intergenerational Transmission of Fertility Patterns." *Oxford Bulletin of Economics and Statistics*, 71(2): 183-208.

Chamberlain Gary and Guido Imbens. 2004. "Random Effects Estimators with Many Instrumental Variables." *Econometrica*, 72(1): 295-306.

Davidoff, Ian and Andrew Leigh. 2008. "How Much do Public Schools Really Cost? Estimating the Relationship between House Prices and School Quality." *Economic Record*, 84(265): 193-206.

Gaines, Brian J. 1998. "The Impersonal Vote? Constituency Service and Incumbency Advantage in British Elections, 1950-92." Legislative Studies Quarterly, 23(2): 167-195.

Ho, Daniel E. and Kosuke Imai. 2008. "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: California Alphabet Lottery, 1978-2002." *Public Opinion Quarterly*, 72(2): 216-240.

Ho, Daniel E., and Kosuke Imai. 2006. "Randomization Inference with Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association*, 101(475): 888-900.

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615-635.

King, Amy and Andrew Leigh. 2009. "Are Ballot Order Effects Heterogeneous?" *Social Science Quarterly*, 90(1): 71-87.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics*, 142(2): 675–697.

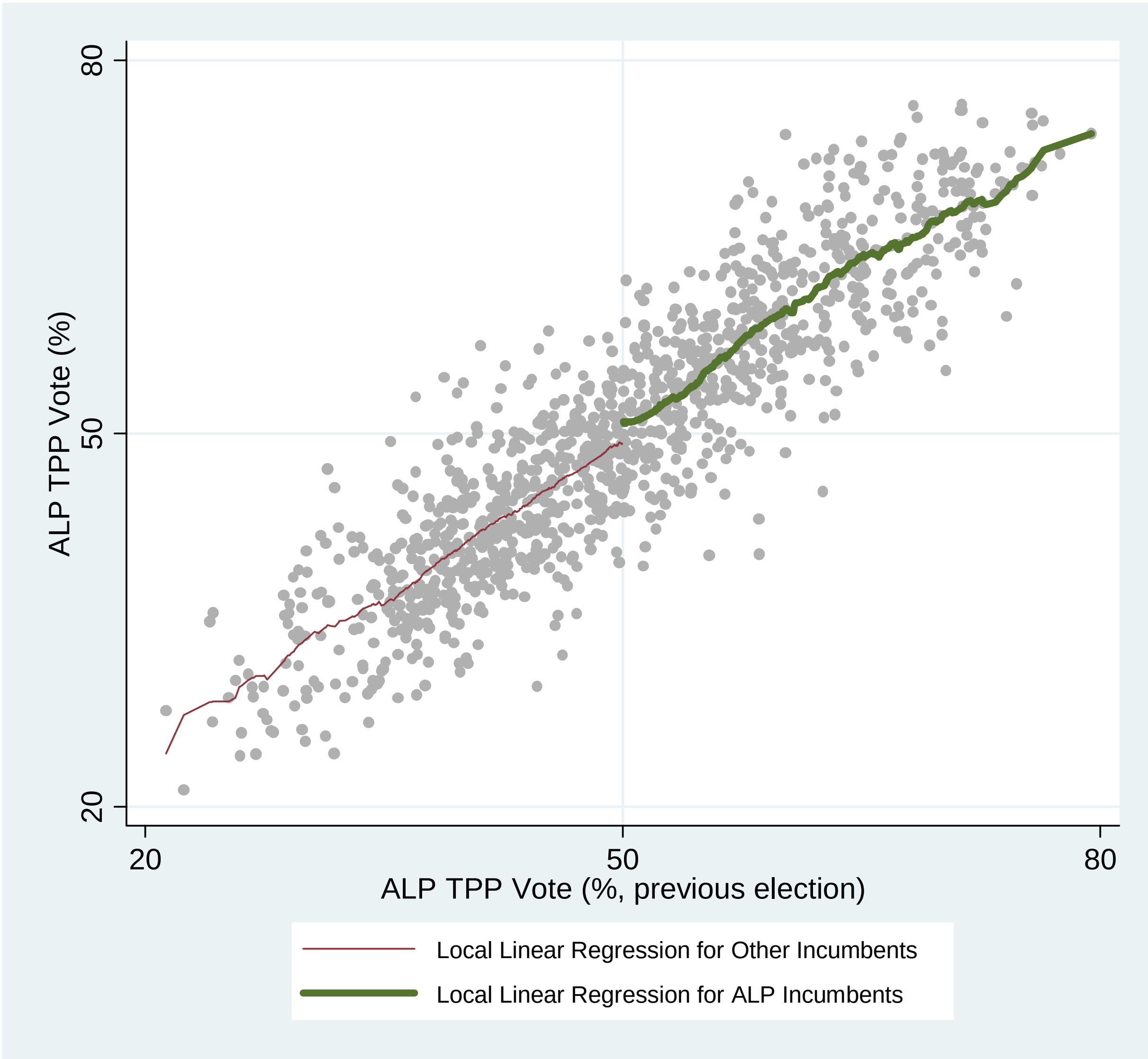
Levitt, Steven D. and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U. S. House." *Legislative Studies Quarterly*, 22(1): 45-60.

Linden, Leigh L. 2004. "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections." MIT, Working Paper.

Middleton, Joel A. and Donald P. Green. 2009. "Do Community-Based Voter Mobilization Campaigns Work Even in Battleground States? Evaluating the Effectiveness of MoveOn's 2004 Outreach Campaign." *Quarterly Journal of Political Science*, 3(1): 63-82.

Uppal, Yogesh. 2009. "The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures." *Public Choice*, 138(1-2): 9-27.

Figure 1: Vote-Share Discontinuity



Note: The number of observations (dots) is 1,141, where observation indicates the two-party preferred (TPP) vote share of the Australian Labor Party (ALP) in the current (1987-2007) and previous (1984-2003) elections in a given electoral division. The lines are drawn based on the local linear smoothing with the rectangle kernel function. The bandwidth, the width of the smoothing window around each point, is a default estimate of STATA 11's **lpoly** command.

Table 1: Vote-Share Discontinuity, Robustness Check

Bandwidth	Default × 50% (1.082)	Default × 100% (2.165)	Default × 150% (3.247)
Local Average Treatment Effect	1.007	1.735	1.175
Standard Error	2.216	1.295	1.136
Bootstrapped Confidence Interval	[-3.445, 5.460]	[-0.867, 4.336]	[-1.107, 3.457]

Note: The number of observations is 1,141. The estimated causal effects are based on the local linear smoothing with the rectangle kernel function, evaluated at the discontinuity (i.e., the ALP vote share in the previous election = 50%). The default bandwidth is an estimate of STATA 11's **lpoly** command. The number of bootstrapped replications is 50, and the level for (normal approximation) confidence intervals is 95%.

Table 2: Boundary Discontinuity

Model	1	2	3	4
Incumbency Dummy (LATE)	9.025	6.350	5.683	5.348
	[0.653]	[0.411]	[0.425]	[0.407]
# of Candidates in Prev. Election = 4	-8.321	-4.814	-3.065	-3.551
	[2.060]	[2.637]	[3.069]	[3.463]
# of Candidates in Prev. Election = 5	-8.667	-4.411	-2.933	-2.492
	[1.919]	[2.412]	[2.864]	[3.343]
# of Candidates in Prev. Election = 6	-8.282	-4.603	-3.201	-2.835
	[2.047]	[2.357]	[2.793]	[3.288]
# of Candidates in Prev. Election = 7	-7.561	-4.385	-3.138	-2.801
	[1.934]	[2.375]	[2.817]	[3.298]
# of Candidates in Prev. Election = 8	-8.179	-4.690	-3.369	-2.926
	[1.901]	[2.384]	[2.830]	[3.291]
# of Candidates in Prev. Election = 9	-9.742	-6.673	-5.280	-5.004
	[1.895]	[2.404]	[2.860]	[3.326]
# of Candidates in Prev. Election = 10	-7.248	-4.059	-3.862	-2.704
	[2.291]	[2.442]	[2.896]	[3.356]
# of Candidates in Prev. Election = 11	-8.228	-5.773	-4.140	-4.252
	[2.366]	[2.671]	[3.082]	[3.553]
# of Candidates in Prev. Election = 12	-3.398	-3.731	-2.785	-3.045
	[2.606]	[2.465]	[3.093]	[3.353]
# of Candidates in Prev. Election = 13	-4.892	-3.41	-3.058	-2.149
	[3.005]	[3.181]	[3.996]	[4.037]
# of Candidates in Prev. Election = 14	-8.421	-6.234		
	[1.863]	[2.381]		
Main Opponent in Prev. Election = LP	4.692	2.048	1.237	0.747
	[1.787]	[2.015]	[2.328]	[2.803]
Main Opponent in Prev. Election = NP	-4.414	-7.028	-1.754	-1.573
	[2.289]	[2.314]	[2.528]	[2.803]
Main Opponent in Prev. Election = GR	-4.489	-6.034	-7.724	-7.735
	[2.278]	[2.453]	[2.713]	[3.152]
Constant	49.099	51.852	49.389	49.714
	[1.513]	[1.366]	[1.637]	[1.818]
R-squared (with-in)	0.300	0.262	0.238	0.234
R-squared (between)	0.283	0.248	0.343	0.340
R-squared (overall)	0.283	0.250	0.281	0.278

Note: All models include polling-place-year fixed effects. The dependent variable is the ALP's two-candidate preferred (TCP) vote share in 1998-2007 elections. Models 1 and 2 include 30,092 observations for 28,710 panels. Models 3 and 4 include 1,449 observations in 729 panels, where the share of total votes within each polling place (in each year) is larger than 10% or smaller than 90%. Clustered robust standard errors are in brackets where clusters are panels.

Table 3: Random Ballot Ordering

	OLS	2SLS	LIML
Incumbency Dummy (LATE)	17.823	17.753	17.631
	[0.407]	[2.772]	[4.542]
# of Candidates in Prev. Election = 3	2.222	2.219	2.213
	[2.485]	[2.468]	[2.475]
# of Candidates in Prev. Election = 4	2.208	2.208	2.208
	[2.433]	[2.412]	[2.412]
# of Candidates in Prev. Election = 5	1.998	2.006	2.020
	[2.437]	[2.436]	[2.469]
# of Candidates in Prev. Election = 6	2.175	2.181	2.192
	[2.449]	[2.439]	[2.460]
# of Candidates in Prev. Election = 7	2.427	2.433	2.444
	[2.470]	[2.461]	[2.483]
# of Candidates in Prev. Election = 8	1.576	1.585	1.602
	[2.490]	[2.496]	[2.544]
# of Candidates in Prev. Election = 9	0.715	0.729	0.752
	[2.555]	[2.587]	[2.678]
# of Candidates in Prev. Election = 10	4.298	4.296	4.292
	[2.667]	[2.645]	[2.648]
# of Candidates in Prev. Election = 11	1.080	1.095	1.122
	[2.931]	[2.966]	[3.068]
# of Candidates in Prev. Election = 12	5.114	5.110	5.103
	[3.527]	[3.499]	[3.505]
# of Candidates in Prev. Election = 13	3.159	3.134	3.090
	[7.174]	[7.179]	[7.294]
Constant	39.481	39.518	39.583
	[2.460]	[2.842]	[3.428]
Partial R-squared of excluded instruments		0.021	
First-stage F-statistic (21, 1102)		1.13	
		(0.305)	
Sagan statistic (over-identification test)		15.277	15.380
		(0.760)	(0.754)

Note: All models include election-year dummies. The dependent variable is the ALP's two-party preferred vote (TPP) share in 1987-2007 elections. The number of observations (electoral divisions) is 1,142. The excluded instruments are ballot-position dummies for ALP candidates and ballot-order dummies for major opponents.