

How Influential is Ballot Design in Elections?*

Agustin Casas[†]
Guillermo Diaz[‡]
Christos Mavridis[§]

October 20, 2020

Abstract

We exploit an original dataset from a referendum in Peru to study the influence of voting “arrangements” on electoral outcomes. The relative importance of these arrangements (e.g., ballot design) with respect to the fundamentals (e.g., ideology, candidates’ quality) has not been measured. After controlling for a comprehensive set of politicians’ characteristics, we estimate unbiased ballot order effects making use of the within party variation in outcomes. We estimate a non-linear probability model and we create counterfactuals to conclude that ballot design not only may have changed the electoral results but also has a greater importance than candidates’ ideology, education, experience and party affiliation.

*We acknowledge the assistance of Fabio Mejias, and the comments from Z. Hessami, M. Machado, R. Pique, P. Riera, J. Millan, I. Ortuño-Ortín, I. Clots, A. Loeper, D. Lougee, and B. Schmidpeter. All errors are ours.

[†]CUNEF and INARBE, acasas@cunef.edu

[‡]Universidad Andrés Bello, jose.diaz@unab.cl

[§]Middlesex University London, c.mavridis@mdx.ac.uk

1 Introduction

Voters usually choose political alternatives from lists on printed ballots. For instance, in the “Australian ballot” (used in most democracies) voters are presented with a list of candidates in which they have to mark their choice. Another example is “ballot propositions”, where voters are presented with a list of proposed reforms, which they can approve or not, one by one.¹ We investigate whether the arrangement of candidates in a piece of paper, i.e., the ballot design, may influence the final outcome of an election (see Wand et al. (2001) for the effect of ballot design on George W. Bush’s 2000 election). In particular, we study how the outcome of forty simultaneous and inter-related elections was affected by the design of the ballot in a developing country.

There is a large body of literature that has focused solely on the effect of the candidates’ placement on the electoral outcomes – i.e., the ballot order effect – and its determinants in developed countries.² We contribute to this literature documenting the ballot order effects in developing countries, and we show that recall elections are unlikely to reflect the opinion of (all) the citizens due to the roll-off produced by the ballot design. Moreover, we study other determinants of voting, such as a politician’s ideology and quality, arguably more substantive issues than ballot design, and we show that they cannot compensate for the documented ballot order effect.

Our data comes from a 2013 recall referendum in Lima, Peru. In this election, citizens were presented with a ballot with the names of the forty municipal council members –or legislators– listed in two columns. For each of these forty names, citizens were asked to indicate whether they wanted the legislator removed or not. The causal interpretation of our results follows from the rules for the allocation of the legislators on the ballot which did not depend on their popularity or other observable characteristics, especially so regarding

¹This system of “ballot propositions” or “ballot measures” is extensively used in California, US. Typically, the propositions imply some sort of legislation or policy reform, presented in the form of referendums.

²Alvarez et al. (2006); Augenblick and Nicholson (2016); Ho and Imai (2006, 2008); Hessami (2016); Chen et al. (2014); Matsusaka (2016) among others. This literature finds that candidates at more prominent locations in the ballot tend to get more votes just because of the ballot design. We call this the ballot order effect.

the column in which they were located. Thus, controlling for precinct unobservables and legislators' observables, we use polling-booth level data to identify the ballot order effect on the share of votes cast to remove, the share of votes cast to keep, and the share of votes not cast or abstentions ("blank" votes). Furthermore, using the legislators' characteristics (political party, experience, age, education, popularity, etc), we quantify the effect of ballot order with respect to these other variables of large political interest in democracies.

We find that the ballot order effect is very large. Being closer to the top of the ballot, as well as being in the first column, increases the share of pro removal votes a legislator gets. In fact the effect of being one step closer to the top of the ballot affects the probability of recall as much as having held a national office in the past. Beyond their position within a column, the column disadvantage is not compensated by any observable characteristic either. In particular, the share of blank votes for legislators positioned in the second column increases by four percentage points, which is around thirty times larger than the effect of having held office in the past. Respectively, the effect on the share of votes for removal is between five and ten times larger, depending on the specification.

In the spirit of King et al. (2000), we also simulate the electoral results under different ballot designs. We show that with a ballot with a random order of legislators, everybody but the mayor would have been removed from office.³ We additionally show evidence that the mechanism behind the ballot order effect is abstention: those who are more prone to abstain as they go down the ballot are those who are against recalling legislators, specially so in regions with a large incidence of poverty.

As a result, the removal referendum, originally designed as a constitutional provision to increase accountability, might have had counterproductive consequences due to a poor implementation that did not take into account (foreseeable) choice fatigue. The twenty legislators listed first in the ballot, representing almost the whole majority coalition, were

³In line with Selb (2008), this counterfactual estimate shows that voters are either prevented from translating their preferences into actual choices or simply cannot lead with the cognitive and/or non-cognitive cost of doing it.

removed and they had to be replaced with a by-election. Hence, the removal referendum may have caused Lima’s council to be less representative of public opinion and voters’ preferences.

2 Institutional Background and Data

Lima, the capital of Peru, has 8.5 million inhabitants and is run by a mayor and thirty nine city legislators.⁴ These positions are chosen in a municipal election every four years by popular vote with a closed-list proportional rule that gives an automatic majority (twenty out of forty legislators) to the party with most votes. All remaining seats are assigned proportionally. Voting is mandatory in Peru and voters’ registries consist of a list of voters in a district, ordered according their “electoral number”. These lists are then split into groups of (up to) 300 voters, following the order in the list.

The Constitutional reform of 1993 added the possibility of calling a recall referendum (*Consulta Popular de Revocatoria*). This provision, meant to keep politicians accountable, implies that all politicians holding office in a sub-national level can be removed from office following a recall referendum. Such recall referendum can take place at any time during a politician’s mandate, except its first and last year and only after a formal request that needs to be signed by 25% of the citizens that live in the jurisdiction of the politician. Up until 2015, there was a cap on the signatures needed, which meant that in the case of Lima only 400 thousand signatures were needed to trigger a recall referendum. Welp (2016) examines the recall referendums in Peru and shows that between 1997 and 2013 there were 5303 recall referendums in 747 out of Peru’s 1645 municipalities.

The electoral rule used in the recall referendums is simple majority rule: if more than 50% of the registered voters participate in the recall referendum and if for a politician the non-confidence votes (**YES**, recall, votes) are more than the confidence ones (**NO**, do not recall, votes), then they are recalled and must be replaced. If less than one third

⁴Unless otherwise noted, we will use the term legislators for all forty politicians, including the mayor.

of the legislators is recalled, parties maintain their number of seats, and the legislators are replaced by substitutes within the party. If more than a third of the legislators is recalled, a by-election is held to replace the recalled legislators.

2010 Election: The 2010 election for the city of Lima’s mayor and 39 members of the municipality’s council was a highly contentious contest between two main coalitions and several other small parties. Susana Villarán, the leader of center-left coalition of parties *Fuerza Social* (FS), was elected mayor. Hence, the municipal council was formed by twenty one legislators from *FS*, thirteen legislators from the coalition between *Partido Popular Cristiano* and *Unidad Nacional* (PPC-UN), two from party *Restauracion Nacional*, and parties *Cambio Radical*, *Somos Peru*, and *Siempre Unidos* had one legislator each.

2013 Referendum: In early 2012, a group of citizens, arguably organized around the figure of the ex-mayor Luis Castañeda, collected signatures to hold a recall referendum. The referendum was set for March 17th 2013 and its ballot design followed these rules: first, parties (or coalitions) were ordered downward according to their number of legislators (the 2010 vote share), and second, within the coalition, legislators were ordered according to their order in their coalition’s (or party’s) 2010 closed list. The forty names were arranged in two columns and the ballot did not include any partisan identification or picture, as shown in Figure 1a. The mayor was located at the top of the ballot. Twenty two legislators ended up being recalled. Since more than a third of the legislators were recalled, a by-election to fill the vacant seats was set for November 2013.

2.1 Data

We use the polling-booth level electoral data of the referendum as well as data on the legislators individual characteristics. Data were provided by *Oficina Nacional de Procesos Electorales* (ONPE) after our request.

There are 36,740 booths divided in 888 centers throughout Lima’s 43 districts. In the referendum the eligible voters number was 6,357,243, and the turnout rate was 83.7%. Figure 2a shows the actual votes for YES and NO for each legislator: the mayor kept her office, while legislators in positions 2 to 21, from *FS*, were recalled. Legislators in position 26 and 31, both from *PPC-UN* were also recalled. The legislator in position 31, Luis Castañeda, was the son of the previous mayor, and he and Villarán were seen as the two competing faces of the referendum. Figure 2b shows the blank vote shares for each legislator.

Table 1 contains a summary of the legislators’ observable characteristics. Variable *Media Exposure* is constructed by counting the mentions of the forty legislators in the period between 2010 and March 2013, in the five most important newspapers in Perú (El Comercio, La República, Perú 21, Gestión and Correo). The median legislator had won a local election once, had no experience in national politics and was almost unknown to the public with only three mentions.⁵

3 Theoretical Background

Most of the literature in ballot structure focuses on two distinct outcomes: roll-off or actual electoral outcomes. The former is defined as a tendency of citizens (who show up to vote) to vote for higher offices and not for the ones that appear lower in a ballot (as in Burnham, 1965). In our paper we estimate the roll-off produced by the ballot order (à la Augenblick and Nicholson, 2016) and then investigate its effect on actual voting outcomes and status quo bias by looking at the removal referendum results (à la Hessami and Resnjanskij, 2019).

Closer to the point in our paper, ballot design is not innocuous as it may favor the candidates ranked first - priming effect (Chen et al., 2014; Esteve-Volart and Bagues, 2012; Meredith and Salant, 2013; Lutz, 2010; Grant, 2017), or ranked last - anti-priming

⁵See Online Appendix for the individual characteristics of each of the 40 legislators.

and/or latency effect (Alvarez et al., 2006). Moreover, the effects of design can be less intuitive in the cases with columns, such as in Blom-Hansen et al. (2016), who show that being placed in the top position in the second column gives the candidates an additional share of four percentage points in Denmark.

In the above mentioned papers, the positioning of the politicians in the ballot is the explanatory variable for voting behavior. That is, individual characteristics, quality and ideology tend to be omitted in the analysis either because alternatives in referendums are not qualified in terms of quality or ideology, or because the politicians' characteristics are bundled in fixed effect estimates. In our setup, we can control for the specific politicians' attributes (ideology or party affiliation, political experience, electoral experience, education, etc).⁶ Hence, we contribute to that literature by weighing the ballot order effect against the more fundamental issues of elections. Moreover, because of the large variance in the legislators' attributes (see Table 1) we inquire whether salient attributes can compensate bad placement, even in a removal referendum where the majority of incumbents are little known, as follows.⁷

Hypothesis 1 (substance): *The ballot design influences voting behavior significantly more than the fundamentals (e.g. party affiliation, experience, etc).*

In line with behavioral research (Grimmelikhuijsen et al., 2017), the explanations of these effects come from psychology and are documented in the political economy literature: the mechanisms pointed out in the literature are related to the voters' cognitive (Shue and Luttmer, 2009) and non-cognitive costs, like fatigue. For instance, Augenblick and Nicholson (2016) suggest the presence of voting fatigue using a natural experiment in California, and they show that when an election appears sooner in a ballot, there is a significantly lower number of abstentions. Consistent with these findings, Ho and Imai (2006, 2008) find that (minor) candidates appearing first in the ballot are more likely to get more votes than if they appear later. Hessami (2016) and Matsusaka (2016) point out

⁶Campbell (2005) uses the term election fundamentals but for U.S. presidential elections.

⁷All the following hypotheses should be read *ceteris paribus*.

that fatigue depends on overloading, i.e., the number of propositions that voters have to decide on. Additionally, Milita (2017) shows that the readability and complexity of the ballot proposals affect individual abstention.

In our case, there are forty legislators listed in two columns and voters should make a decision (remove or not) for each of them, i.e., they face forty simultaneous elections. Hence, to begin with, we are in the presence of overloading, which is ubiquitous in ballot order studies (Alvarez et al., 2006). Additionally, the implications of the referendum depend on how many legislators are removed: if less than a third are removed, then they are replaced with the substitutes, but if more than a third are removed, then they are replaced with a by-election. Hence, the rules of this referendum (or 40 referendums) are extremely complex, relative to the literature.

In accordance to the idea that “fatigue” would increase roll-off as voters move down the ballot, the order of the legislators within a column should influence voting behavior. That is, the options presented last should see more blank votes.

Hypothesis 2 (roll-off): *Within a column, the options ranked last suffer from more blank votes due to roll-off.*

Related to the point above, the introduction of columns makes the ordering less straightforward. It could be the case that both legislators ranked first in each column are subject to priming rather than the first one in the first column. In this election, each legislator is accompanied by a number that emphasizes a particular order: 1 to 20 for the first column, and 21 to 40 for the second. Thus, the first one in the second column is the 21st election.

We then conjecture that the order that dominates priming effects (and roll-off) is determined by the numbering of legislators shown in Figure 1a. In this regard the first column is primed, which leads us to formulate the following hypothesis:

Hypothesis 3 (column): *The second column is more affected by blank votes than the first one. Moreover, the 20th option (last of first column) has more blank votes than the*

21st (first of second column) due to the jump of column, numbering and roll-off.

As is common in studies of ballot design – and elsewhere– citizens are unlikely to be well informed of all alternatives. In this case, with a long and complex ballot, in a developing country, the lack of information may be exacerbated in poorer regions. And so, roll-off may “kick in” earlier in the ballot.

Additionally, in this election, there are strategic considerations. Since *FS* –the left-wing coalition– had an absolute majority in the city council, left-wing voters are more interested in keeping the seat distribution and so have less incentives to remove legislators. Left- and right-wing voters’ anticipation of these considerations may be intertwined not only with political preferences, but also with cognitive and non-cognitive costs. Furthermore, when voters lack information that is needed to decide due to the environment or individual circumstances (Brockington, 2003; Miller and Krosnick, 1998), behavioral cues are more prevalent. Shue and Luttmer (2009) show that this is the case where cognitive costs are higher, with an emphasis in poorly educated and poor regions. In the same spirit, we pose that the prevalence of this behavior may be higher in districts with higher incidence of poverty, which coincidentally is where *FS* performed better (c.f. district Villa El Salvador).

Hypothesis 4 (poor): *The prevalence of roll-off due to ballot design is greater in poorer districts.*

Therefore, if the voting behavior of left-wing supporters is more likely to be affected by roll-off then these voters are less likely to vote for alternatives in the second column. Thus, the politicians placed in that column, who happen to be mostly *PPC* and other right-wing and minor parties, are more likely to obtain fewer **YES** and **NO** votes. As mentioned earlier, the overlap between *FS* and poverty is high, i.e., the correlation between the share of *FS* and the share of poor people across districts is 0.70. Thus, the coalition and parties in the second column – especially the right-wing coalition *PPC*– would benefit

from a larger decrease in YES votes than in NO votes. On the contrary, the politicians placed in the first column are more likely to be harmed by the ballot design.

Hypothesis 5 (FS vs PPC): *The ballot design influences the outcome of the election, benefiting Partido Popular Cristiano (PPC) at the expense of Fuerza Social (FS).*

In sum, if the hypotheses are confirmed, ballot order affects voting behavior, in particular, it increases roll-off. The voters who stop voting as they go down the ballot or – especially – as they change columns are more likely to be *FS* supporters and/or from poorer districts. Thus, the design benefited *PPC*. Moreover, even though some legislators were ex-ante better politicians than others, this quality/ideology advantage does not compensate the ballot order effects.⁸

4 Empirical strategy

We study the effect of legislators’ ballot location on their performance in the recall election. To do this, we relate the share of votes for (YES) or against (NO) the removal of legislator with their in the ballot. Both legislators and their respective locations are indexed by l . We aggregate electoral results at each polling center indexed by i , which are located in precincts (districts), d . We estimate variants of the following specification:

$$Y_{idl} = \gamma' \mathbf{Z}_l + \beta' \mathbf{X}_l + \alpha' \theta_d + \varepsilon_{idl} \quad (1)$$

where Y_{idl} is the outcome (share of YES or NO votes) of legislator l in polling center i (located in precinct d), \mathbf{Z}_l is a vector of variables that identify the location of legislator l in the ballot, and \mathbf{X}_l is a vector of legislators’ characteristics, including their party’s share in 2010 elections. We also include fixed effects for each precinct (district), θ_d , to account for heterogeneity.

⁸As such, the paper is also related to the literature on political institutions in developing countries and their abuse (see for example Gutiérrez-Romero (2013) and references therein).

We are interested in estimating the vector of parameters γ that capture the effect of location in the ballot. The most important variables in \mathbf{Z}_l are $Order_l$ and $Column_l$. The former, $Order_l$, takes values from 1 to 40, for each of the legislators' consecutive locations. The latter, $Column_l$, takes value 1 if the legislator is located in the second column of the ballot, and zero otherwise. We also include the interaction $Order_l \times Column_l$ to allow for different effects of $Order$ in each column.

In all specifications we cluster the standard errors by official and polling center ($Center \times Order$), and in all estimations of Equation 1, we include district fixed effects. Lastly, among the controls we include coalition fixed effects, hence we can interpret our results as within-party order effects.

Identification strategy. The main threat to the identification of γ is the possible endogeneity of \mathbf{Z}_l , i.e., correlation between the legislators' order or column location with the unobserved preferences for or against their recall, ε_{idl} . We rely on features of the institutional setting and appropriate controls in \mathbf{X}_l (and the fixed effects) to show that this correlation is unlikely in our model. In other words, the causal interpretation of our results follows from the rules for the allocation of the legislators on the ballot, especially the identification of the column effect, and the granular data that limits the influence of omitted variables.

Regarding the order of legislators **within** a party, the subject of Hypothesis 2, notice that in 2010 the two most important lists, *FS* and *PPC-UN* (which obtained 35 legislators, including the mayor) were both coalitions of parties, elected by proportional rule with a closed-list system. In closed-list elections voters do not indicate a preference for individual candidates, but for lists; hence candidates located near the top are more likely to be elected to the city council. Then, the order of candidates *within* the coalitions' closed lists likely depends on the parties' bargaining power within the coalition, and not necessarily on personal characteristics, skill, or performance in office, as we show in the

following section.⁹ By including dummies for the major coalitions in the vector of control variables, \mathbf{X}_i , we exploit this within-party variation to estimate the effect of *Order* on electoral performance.

Additionally, whether the legislators are located in the first or second column – the subject of Hypothesis 3– also depended on the design of the ballot paper, i.e., on the order **across** parties. The design was decided by the *Oficina Nacional de Procesos Electorales* (ONPE) only after the recall referendum was accepted. ONPE is an autonomous institution with maximum authority in electoral matters in Peru. Its authority to organize and oversee elections is warranted by Peru’s Constitution, and its president is elected by an independent institution that belongs to the judiciary branch (called, *Consejo Nacional de la Magistratura*). That is, neither the referendum nor the ballot design could have been anticipated and so, the placement across columns is plausibly exogenous, specially so for the *FS* legislators.

Finally, to address potential confounding factors due to correlation between elections (2010 and 2013 preferences), we control for the electoral results of the coalition in 2010 in the district where the polling center is located. The name of this variable is *Party Share in 2010 elections*. Given that we also have district-level fixed effects, this allows us to separate the effect of preferences from location in the ballot to estimate the effect of *Column* location on electoral performance.¹⁰

4.1 Validity of the identification strategy

In the online appendix we confirm the robustness of our identification strategy using an instrument for the legislators’ ballot order and we address a number of potential

⁹Figure 1b shows the arrangement of the coalitions, identifying each legislator’s party membership. The order of legislators in the largest coalition illustrates our point about the bargaining between the parties in the coalition. See, for example, that the *FS* legislators in positions 3, 8, 15 and 22, belonged to party *Movimiento Tierra y Libertad*, and that between them there were legislators coming from other parties.

¹⁰Notice that while the fixed effects control for all possible district-level confounding variables, observed and unobserved, the *Party Share in 2010 elections* captures the district level support to each party, separately.

confounding factors. In particular, we show that (i) political parties did not select the order of legislators in 2010 based on candidates' observable measures, (ii) if parties had selected the order based on unobservable characteristics, we would be under-estimating our coefficients, (iii), voters have little information about individual legislators, and (iv) the voters' political preferences in a polling booth are independent of the legislators' order.

5 Results

A careful inspection of figures 2a and 2b shows qualitative evidence in support of H2 and H3 (and partly H5). Figure 2a illustrates that while the number of YES votes (i.e., to remove politicians) decreases continuously within the first column, the jump to the second column reflects a discontinuous decrease, even up to a point in which the NO overtakes the YES (i.e., against the removal of legislators).

A similar pattern emerges in Figure 2b: roll-off increases the blank votes continuously within columns (H2), except for the jump between the legislators numbered 20 and 21. Thus, the columns' jump is reflected in a jump in blank votes as well (H3), i.e., many people stop voting. We quantify these effects (and others) below using the specification in equation 1.

5.1 Analysis of the probability of being removed from office

Table 2 shows the results of estimating equation 1 taking the shares of YES and NO as dependent variables. The basic findings of columns (2) and (3) confirm what Figure 2a shows: going down one position in the ballot decreases the share of YES and NO votes asymmetrically. Similarly there is a strong effect of being in the second column, as the shares of YES and NO votes both decrease significantly: being in the second column decreases the YES votes by 2.6 percentage points (at a decreasing rate due to the positive interaction with order). These effects are not entirely explained by ideological positioning

of the voters as we control for the *Party Share in 2010*. The higher the district share of the party the legislator belonged to, the lower the share of YES votes. The opposite is true for the NO votes. In particular, a 1% increase in the votes for the legislators' party implies 1.3 fewer percentage points of YES votes, and 4.4 percentage points of voting against the recall. This suggests that a district's ideological leaning may have had an effect on the probability of recall of a legislator. However, other legislator characteristics, such as gender, whether they are College graduates or they had experience in politics seem to have very little effect on the shares: they are either extremely small in magnitude or not statistically significant. This is in line with our assumption that voters had limited information on the legislators and their individual characteristics.

In a large city like Lima, the district heterogeneity might have some effect on the electoral results. We control for these demographic characteristics by including an interaction of poverty rate (the share of citizens living under the poverty line) with *Order*, *Column* and their interaction. Columns (5) and (6) show these results. The poverty rate strengthens the effect of *Order* and *Column* for the votes for the recall of legislators: the greater the poverty, the more the share of YES decreases as we go down the ballot. The effect on NO votes is the opposite.

In the comparison between the differential effects on YES and NO in poorer districts, it is worth looking at the effects of order with and without the poverty interaction. The effect of order in the first twenty positions (the left column of the ballot) makes more likely to recall worse positioned legislators (i.e., $-0.00069 < -0.00060$ in columns (2) and (3)). However, columns (5) and (6) show that this is reversed in richer neighborhoods (i.e., $-0.00035 > -0.00060$). In poorer districts, looking at the sum of the coefficients of order and its interaction with poverty, YES votes decrease at a rate four times larger than the NO votes (i.e., $-0.00246 < -0.00059$).¹¹

This large heterogeneity is also observed in the jump between the left and right columns, up to a point in which the column effect is reversed. That is, while in richer

¹¹All sums and differences are statistically significant at the 95% level.

districts the column effect is pro recall –the NO share decreases faster than the YES share– in poorer ones it is the opposite –keeping parties fixed. The heterogeneous effect of being placed in the second column depends greatly on poverty: in poorer regions the effect of column on YES is almost six times larger than in richer ones (in absolute value). Similarly, the votes against removal are 9 percentage points larger in poorer districts. While in richer neighborhoods being in the second column implies 2.6 fewer percentage points, in the poorer ones NO for the legislators placed in the second column are 5.8 percentage points larger than for those placed in the first one.

Analysis by district. Table 3 shows the results of the least squares regression in two districts: Villa El Salvador and San Isidro. Villa El Salvador is the district where *FS* obtained their highest share in the 2010 election, whereas San Isidro is the district where *PPC-UN* got their highest share, which is the same district where *FS* got their lowest share. Although all the coefficients are consistent with our previous results, some very interesting patterns emerge. The effects of *Order* and *Column* are larger – in absolute terms– in a left-wing district than in a right-wing one. Hence, the design of the ballot seems to have more pervasive effects in the district where the mayor was elected with the largest share.

5.1.1 Counterfactual Analysis

To analyze the probability of YES and NO for each individual legislator more concretely we also estimate a multinomial probit and we study what would have happened in Lima if the design of the ballots were different.¹² We follow two different strategies: (i) we calculate the predicted values, eliminating the order and column effects, and (ii) we simulate the results using a ballot with a random order of legislators.¹³

(i) In order to see how the ballot design affected the referendum result we force the

¹²We only report the counterfactuals in the current version, for economy of space. The full results can be provided immediately upon request.

¹³We perform the analysis by district for the 43 districts, we collect the estimated coefficients and, taking into account the districts size, we perform our counterfactual exercises.

Column and *Order* effects to be zero for all officials, by setting the variables to zero. Hence we obtain the predicted values as if the ballot design had no effect, i.e., the “no order” counterfactual: Figure 3a shows this result. We see that everybody except the mayor would have been recalled.

In practical terms, this result shows what the outcome would have been if every voter received a different single-column ballot where the order of the officials was completely randomized.

(ii) We also simulate the election by randomizing one thousand times the order of the forty legislators in a single column. In almost all repetitions, it is only the mayor who is not getting recalled.¹⁴ More interestingly, we also investigate what would happen with two columns, where the order of parties is random as well as the order of the legislators within the party. Figure 3b shows that it is always the case that some legislators from the largest coalition will end up in the first column. Hence, they are all equally likely to be recalled.¹⁵

The predicted values and the simulation point into the same direction: the ballot design gave an “unfair” advantage to all the legislators located in the second column. A ballot randomization would have ended up in a completely different electoral outcome.

After the by-election the final composition of the legislature included seven new *PPC* legislators, which added to the original ten *PPC-UN* legislators elected in the 2010 election, increased the party’s seats to 17, even though the share of the party between the two elections had actually decreased.¹⁶ In fact, the only way for *FS* to return to their original number of seats (21 out of 39, excluding the mayor), would have been to get 91% in the by-election. This fictional percentage would show an extremely high support for

¹⁴Figure A2 in Online Appendix.

¹⁵See the online appendix for the details of the simulations.

¹⁶*PPC* vote share in 2010: 37.6%. *PPC* vote share in 2013: 29.6%. It should be noted that in 2010 the *PPC-UN* coalition included party *Perù Possible* and one of the legislators elected with *PPC-UN* came from *Perù Possible*. However, in 2013 *Perù Possible* ran on its own and got 11%. Since *Perù Possible* ran on its own in 2013 we do not include its single legislator elected with *PPC-UN* in *PPC-UN*’s 2013 numbers.

the coalition that would barely control the legislature.

5.2 Suggested Mechanism

What is the mechanism behind the documented ballot order effects? In this section we show suggestive evidence that (i) ballot design affected blank voting, and that (ii) blank voting affects asymmetrically the vote for or against the recall. In particular, it seems that, keeping unobserved preferences fixed, blank votes increase at the expense of votes that would have been against the recall, especially for legislators in the first column of the ballot.

Determinants of voting blank: As shown in Figure 2b, the position of the legislator on the ballot has a very strong effect on his share of blank votes: in Table 2 the coefficients of *Order*, *Column* and their interaction are all statistically significant at the 1% level. While *Order* and *Column* are positive, indicating that being further down the ballot leads to more blank votes, the interaction coefficient is negative. Being ten positions further away in the first column implies an increase in blank votes by 1.3 percentage points. Furthermore, being in the second column increases the blank votes by more than 4 percentage points. Note that we included dummies for the Mayor and for Luis Castañeda. In the second column the blank share is increasing at a lower rate compared to the first one (ten positions further down in the second column implies an increase in blank votes by 0.0053 percentage points). One possible interpretation of this result is that voters who keep voting in the second column may have a lower cost of voting. This interpretation may have implications beyond the decision to abstain: if voters “self-select” into the second column, their preferences and their voting behavior may differ in comparison to the set of voters in the first column. We correct this issue by controlling for voters’ preferences in the previous election (variable *Party share in 2010*). These results are robust to including poverty interactions at the district level.

Blank voting induced sample selection in the referendum: Our data allows us to compare polling booths within a polling center, keeping all potential unobservables fixed, to gauge the correlation between blank voting and the preferences for or against the recall of each legislator

Thus, we compare booths with different shares of blank votes for a legislator and see if there is a correlation with the share of YES votes, in favor of the recall of the legislator. The latter share is expressed as a fraction of the total YES and NO votes (non-blank votes), and it is then an indicator of preferences conditional on not having voted blank.

We plot the coefficients in Figure 4, with their corresponding 95% confidence intervals. We can see that there is a positive statistically significant effect (at the 95% confidence level) for almost every ballot location. For instance, a 10% increase of blank votes is associated with a 14% increase in the votes for recall for Susana Villaran (placed first). This figure shows that blank votes have a weaker impact on the share of YES for those who appear later in the ballot. While the average effect in the first column is 0.075 the second column is around 0.03. This suggests that blank-voting effect on the votes for recall was stronger for legislators in the first column. Given that we see a monotonic increase of blank voting, this also means that the subpopulation that stopped voting between columns created biased results by not expressing their views on the second column.

5.3 Discussion

One of our main contributions is the quantification of the effect of order – especially column– in comparison to the most salient attributes of candidates in democracies. The effect of order and column is not only statistically significant and very robust, but it is also significant and very large in political and economic terms. For instance, a comparison with our measures of political experience are quite illustrative: Column (1) in Table 2 shows that the effect of order (being one step further from the top) on blank votes is – with opposite sign – almost identical in magnitude to having won a national election beforehand. With respect to column, all measures of political experience, either competing

or winning an election, do not get even close to compensate 10% of the column effect on blank votes.

Another very interesting comparison comes from ideology, loosely measured by party affiliation. In the same table, the effect of column on share of blank votes is –at the very least– ten times larger than the effect of being affiliated to the right-wing coalition (*PPC-UN*), with respect to the left-wing coalition (*FS*). On the contrary, the effect of *Order* is three or four times smaller: being affiliated to *PPC-UN* would have assured a number of abstentions comparable to a politician affiliated with *FS* who appears three positions closer to the top in the ballot. If we look at the effects of poverty we obtain larger coefficients (the one associated with $Order \times Poverty$ is 0.0022).

To sum up, while the explanatory power of our model does not change much by including the interaction with some sociodemographic characteristics, we see that the effect of order seems to depend on them: in poorer districts the effect of ballot design increases. Hence, it is not only that the instruments of voting influence the election more than the officials (i.e., their ideology and quality), but, even more worrisome, they affect it even more in the disadvantaged districts.

These results hinge on Lima’s institutional setting. For instance, mandatory voting implies that the selection of voters is different than the voters under voluntary voting. According to the Institute for Democratic and Electoral Assistance, thirty countries currently practice compulsory voting. Thus, even though there could be external validity concerns in the case of developed countries with voluntary voting –where the effect could be argued to be smaller– our results are a good estimate for countries, regions and cities with mandatory voting, like South America. A different issue is the external validity in a context with perfect information about the candidates/alternatives, however –as mentioned in Section 3– when ballots are not especially long or complex, the order effects are *per se* less pervasive.

6 Conclusion

On one hand, Peru’s constitution allows for an ad-hoc recall of politicians during their tenure. This provision, together with regular elections, makes politicians more accountable to their citizenry. On the other hand, we showed that the implementation of this provision and the “removal referendums” should take into account electoral details –such as ballot design– to ensure that preferences are translated into electoral results. That is, a poorly designed ballot may have consequences that go against the spirit of the law.

Ultimately, one would like to address whether such constitutional provision reinforces the removal of low-performing incumbents and at the same time ensures that high-performing ones would keep their offices. Nonetheless, until the last reform of removal referendums in Peru, this channel of accountability has been distorted by the losing parties. In our study, the opposition announced the intention to hold a removal referendum only a few months after *FS* took office, impeding the course of a regular legislative term.¹⁷

Our results indicate that citizens vote following the order in which legislators are listed in the ballot, respecting the order of the columns: they sequentially make a decision for each legislator in the first column, and then they move to the second one. They do so, irrespective of the ideology (political party) and quality (education, experience, etc) of the politicians under scrutiny. Hence, as voters move down the ballot, they are more likely to abstain, specially in disadvantaged locations.

Unfortunately, our case of study shows yet another instance in which institutional design has unintended (but foreseeable) consequences. Institutional reform has to be accompanied with a thoroughly thought implementation. Electoral transparency scholars and international electoral observers may also need to devote attention to ballot design. If these “details”, like ballot design, may tilt elections in developed and strong democracies, younger democracies’ accountability mechanisms may be threaten by a greater exposure

¹⁷We thank an anonymous referee for highlighting this issue. Additionally, there is a lack of data that prevents us from analyzing the short-spanned untainted performance of the incumbents.

to manipulation.

7 Tables and Figures

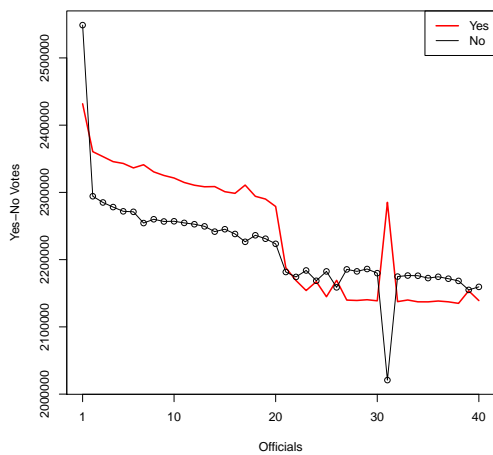


(a) Ballot as seen by the voters

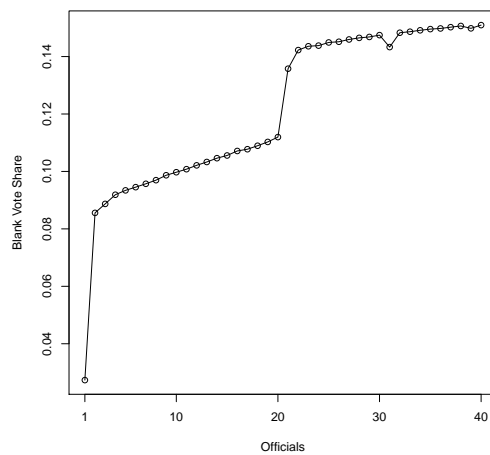


(b) Ballot as published in a newspaper

Figure 1: Ballot for referendum

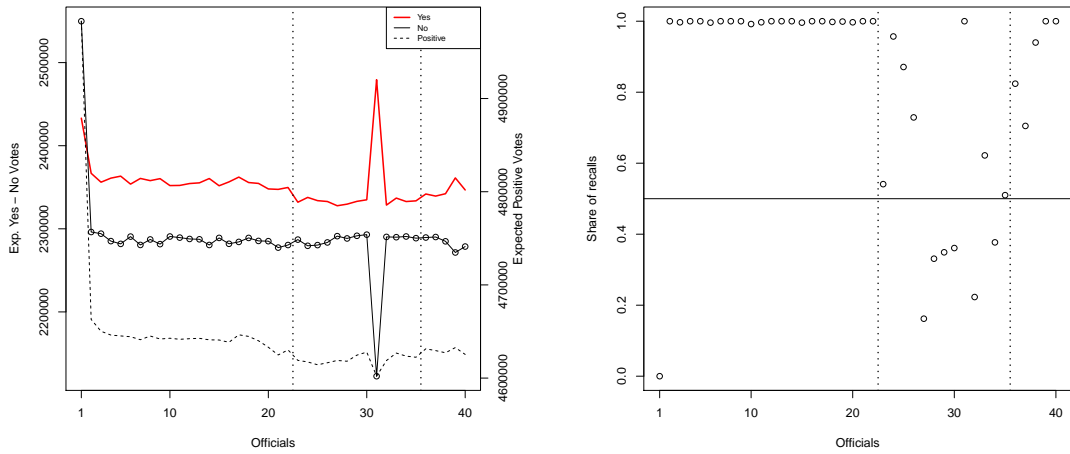


(a) Yes (red) and No votes



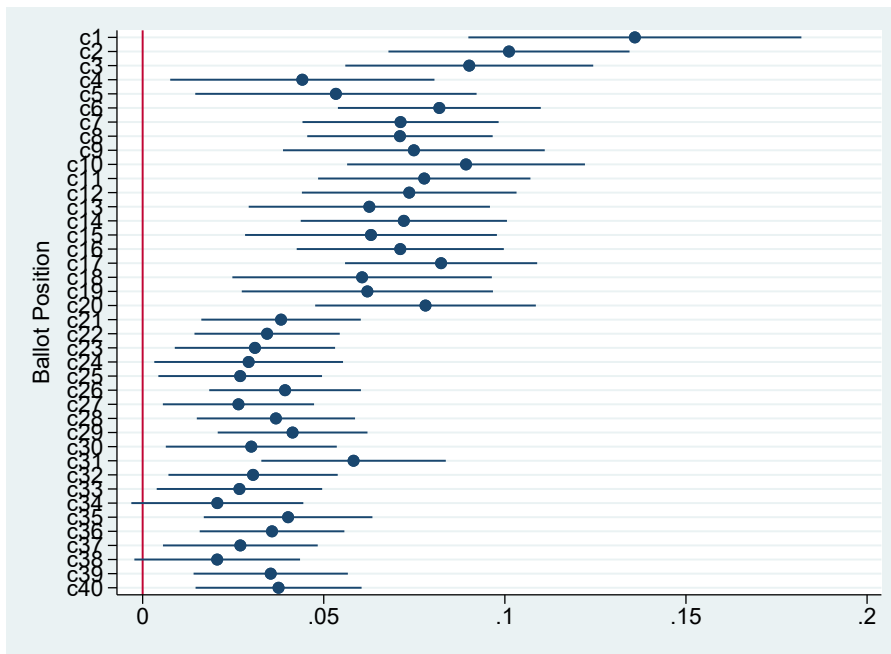
(b) Share of blank votes

Figure 2: Referendum



(a) Predicted values: no column, no order (b) The vertical axis indicates the share of simulations in which each official (ordered from 1 to 40) would have been removed given ballots with two columns and random order of parties and random order of legislators within parties across districts.

Figure 3: Counterfactuals



The figure plots the coefficient and confidence intervals of a regression of YES/YES + NO on blank votes, for each of the forty officials.

Figure 4: Effect of Blank votes on recall votes for each legislator

	Minimum	Maximum	Average	Median	Sd
Age	26	70	46.95	49	11.97
Gender	0	1	0.675	1	0.47
Times Candidate National	0	3	0.42	0	0.81
Times Candidate Local	1	5	1.57	1	0.98
Times Elected Local	1	5	1.27	1	0.78
Imputed Years of Education	11	20	15.47	15.5	1.94
Media Exposure	0	1068	42.35	3	171.3

Gender is a variable that takes value 1 if the official is male. Times Candidate (National and Local) count the number of times the politician was a candidate in national and local elections. Times elected local counts the local elections won by the politician. The years of education have been imputed using the highest known educational level of the official, taking half of the duration if this educational level was not completed. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 1: Summary Statistics of Legislators

	(1)	(2)	(3)	(4)	(5)	(6)
	Blank Share	Yes Share	No Share	Blank Share	Yes Share	No Share
Order	0.0013*** (0.0000)	-0.0007*** (0.0000)	-0.0006*** (0.0000)	0.0010*** (0.0000)	-0.0004*** (0.0001)	-0.0006*** (0.0001)
Column	0.042*** (0.0009)	-0.026*** (0.0012)	-0.017*** (0.0011)	0.039*** (0.0015)	-0.014*** (0.0025)	-0.026*** (0.0024)
Column × Order	-0.0008*** (0.0000)	0.0003*** (0.0001)	0.0005*** (0.0001)	-0.0006*** (0.0000)	-0.0002** (0.0001)	0.0009*** (0.0001)
Mayor	-0.033*** (0.0056)	0.0030 (0.0063)	0.030*** (0.0060)	-0.034*** (0.0046)	0.0031 (0.0059)	0.030*** (0.0060)
Castaneda	0.0022* (0.0012)	0.031*** (0.0039)	-0.033*** (0.0036)	0.0021** (0.0010)	0.031*** (0.0039)	-0.033*** (0.0037)
Right Party	0.0044*** (0.0006)	-0.0049*** (0.0006)	0.0004 (0.0006)	0.0044*** (0.0005)	-0.0048*** (0.0006)	0.0004 (0.0006)
Minor Parties	-0.0052*** (0.0009)	-0.0068*** (0.0014)	0.012*** (0.0014)	-0.0017** (0.0008)	-0.0082*** (0.0015)	0.0100*** (0.0014)
Party Share in 2010 elections	-0.031*** (0.0010)	-0.013*** (0.0036)	0.044*** (0.0033)	-0.018*** (0.0015)	-0.018*** (0.0043)	0.037*** (0.0039)
College	-0.0008*** (0.0002)	0.0003 (0.0003)	0.0004* (0.0003)	-0.0007*** (0.0002)	0.00033 (0.0003)	0.0004 (0.0003)
Age	0.0000 (0.0000)	0.0000 (0.0000)	-0.0000** (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	-0.0000** (0.0000)
Gender	0.0002 (0.0001)	0.0007*** (0.0003)	-0.0009*** (0.0002)	0.0001 (0.0001)	0.0007*** (0.0003)	-0.0008*** (0.0002)
Candidate in national elections	-0.0000 (0.0001)	0.0001 (0.00020)	-0.0001 (0.0002)	-0.0000 (0.0001)	0.0001 (0.0002)	-0.0000 (0.0002)
Candidate in local elections	-0.0000 (0.0001)	0.00033 (0.0002)	-0.0003 (0.0003)	0.0000 (0.0001)	0.0003 (0.0002)	-0.0003 (0.0002)
Elected in national elections	-0.0015* (0.0009)	0.0029*** (0.0009)	-0.0013 (0.0013)	-0.0012** (0.0005)	0.0027*** (0.0009)	-0.0015 (0.0011)
Elected in local elections	-0.00022 (0.0002)	-0.0004 (0.0003)	0.0006** (0.0003)	-0.0002 (0.0001)	-0.0004 (0.0003)	0.0006** (0.0003)
Media Exposure	-0.0000*** (0.0000)	0.0000 (0.0000)	0.0000*** (0.0000)	-0.0000*** (0.0000)	0.0000* (0.0000)	0.0000*** (0.0000)
Order × Poverty				0.0022*** (0.0002)	-0.0021*** (0.0004)	-0.0001 (0.0003)
Column × Poverty				0.017* (0.0089)	-0.075*** (0.012)	0.058*** (0.011)
Column × Order × Poverty				-0.0009** (0.0004)	0.0034*** (0.0005)	-0.0025*** (0.0005)
Constant	0.100*** (0.0006)	0.46*** (0.0014)	0.44*** (0.0014)	0.096*** (0.0006)	0.47*** (0.0016)	0.44*** (0.0015)
N	35520	35520	35520	35520	35520	35520
R ²	0.83	0.90	0.93	0.83	0.90	0.93

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

District fixed effects. All regressions are clustered at the Center × Order level. Minor Parties is a dummy that takes value 1 if the official belonged neither to the party of the Mayor neither to the right-wing coalition. Party share in 2010 elections is the share that the official's party obtained in the 2010 elections in a district. College is a dummy variable that takes value 1 if the official had at least finished college. Gender is a dummy taking the value 1 if the official is a man. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election. All numbers rounded to keep only four decimal places.

Table 2: Least Squares estimations of Blank, Yes and No shares equations (school level).

	Villa El Salvador		San Isidro	
	(1) Yes Share	(2) No Share	(3) Yes Share	(4) No Share
Order	-0.0009*** (0.0000)	-0.0008*** (0.0000)	-0.0007*** (0.0001)	-0.0002 (0.0001)
Column	-0.032*** (0.0013)	-0.017*** (0.0010)	-0.017*** (0.0056)	-0.013** (0.0054)
Column \times Order	0.0006*** (0.0001)	0.0004*** (0.0000)	-0.0002 (0.0003)	0.0007** (0.0003)
Mayor	0.013 (0.0092)	0.035*** (0.0074)	-0.0083 (0.027)	-0.0013 (0.026)
N	1080	1080	320	320
R^2	0.98	0.98	0.97	0.94

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

School fixed effects. Robust standard errors. Controls are not reported but available upon request (Castaneda, Minor Parties, College, Gender, Candidate variables, and Media Exposure). All numbers rounded to keep only four decimal places.

Table 3: Least Squares estimations of Yes and No shares equations for two districts (school level)

References

- Alvarez, R. M., B. Sinclair, and R. L. Hasen (2006). How Much Is Enough? The “Ballot Order Effect” and the Use of Social Science Research in Election Law Disputes. *Election Law Journal* 5(1), 40–56.
- Augenblick, N. and S. Nicholson (2016). Ballot position, choice fatigue, and voter behaviour. *Review of Economic Studies* 83(2), 460–480.
- Blom-Hansen, J., J. Elklit, S. Serritzlew, and L. Riis Villadsen (2016). Ballot position and election results: Evidence from a natural experiment. *Electoral Studies* 44, 172–183.
- Brockington, D. (2003). A low information theory of ballot position effect. *Political Behavior* 25(1), 1–27.
- Burnham, W. D. (1965). The Changing Shape of the American Political Universe. *American Political Science Review* 59(1), 7–28.
- Campbell, J. E. (2005). The fundamentals in US presidential elections: Public opinion, the economy and incumbency in the 2004 presidential election. *Journal of Elections, Public Opinion and Parties* 15(1), 73–83.
- Chen, E., G. Simonovits, J. A. Krosnick, and J. Pasek (2014). The impact of candidate name order on election outcomes in North Dakota. *Electoral Studies* 35, 115–122.
- Esteve-Volart, B. and M. Bagues (2012). Are women pawns in the political game? Evidence from elections to the Spanish Senate. *Journal of Public Economics* 96(3-4), 387–399.
- Grant, D. (2017). The ballot order effect is huge: evidence from Texas. *Public Choice* 172(3-4), 421–442.
- Grimmelikhuijsen, S., S. Gilke, A. L. Olsen, and L. Tummers (2017, jan). Behavioral Pub-

- lic Administration: Combining Insights from Public Administration and Psychology. *Public Administration Review* 77(1), 45–56.
- Gutiérrez-Romero, R. (2013). Decentralisation, Accountability and the 2007 MP Elections in Kenya. *Journal of Development Studies* 49(1), 72–94.
- Hessami, Z. (2016). How Do Voters React to Complex Choices in a Direct Democracy? Evidence from Switzerland. *Kyklos* 69(2), 263–293.
- Hessami, Z. and S. Resnjanskij (2019). Complex ballot propositions, individual voting behavior, and status quo bias. *European Journal of Political Economy* 58, 82–101.
- Ho, D. E. and K. Imai (2006). Randomization Inference With Natural Experiments. *Journal of the American Statistical Association* 101(475), 888–900.
- Ho, D. E. and K. Imai (2008). Estimating causal effects of ballot order from a randomized natural experiment: The California alphabet lottery, 1978-2002. *Public Opinion Quarterly* 72(2), 216–240.
- King, G., M. Tomz, and J. Wittenberg (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American journal of political science*.
- Lutz, G. (2010). First come, first served: The effect of ballot position on electoral success in open ballot PR elections. *Representation* 46(2), 167–181.
- Matsusaka, J. G. (2016). Ballot order effects in direct democracy elections. *Public Choice* 167(3-4), 257–276.
- Meredith, M. and Y. Salant (2013). On the Causes and Consequences of Ballot Order Effects. *Political Behavior* 35(1), 175–197.
- Milita, K. (2017). Beyond roll-off: individual-level abstention on ballot measure voting. *Journal of Elections, Public Opinion and Parties* 27(4), 448–465.

- Miller, J. M. and J. Krosnick (1998). The Impact of Name Order on Election Outcomes. *Public Opinion Quarterly* 62(3), 291–330.
- Selb, P. (2008). Supersized votes: ballot length, uncertainty, and choice in direct legislation elections. *Public Choice* 135(3-4), 319–336.
- Shue, K. and E. F. P. Luttmer (2009). Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes. *American Economic Journal: Economic Policy* 1(1), 229–257.
- Wand, J. N., K. W. Shotts, J. S. Sekhon, W. R. Mebane, M. C. Herron, and H. E. Brady (2001). The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. *The American Political Science Review* 95(4), 793–810.
- Welp, Y. (2016). Recall referendums in Peruvian municipalities: a political weapon for bad losers or an instrument of accountability? *Democratization* 23(7), 1162–1179.