Learning-By-Losing: Do Political Parties Widen Representation to Win Elections?[*]

Su-Min Lee[†]

University of Cambridge

November 10, 2020

Abstract

Although the right to vote has long been near-universal in many countries, there are still significant inequalities in political representation. One contributing factor may be political parties selecting candidates to try and win elections. I argue that parties may learn to be more inclusive in their selection in the aftermath of a defeat, through a 'learning-by-losing' process. The case study of the Conservative Party after the shock landslide defeat in the UK 1906 election provide a unique setting to identify the response of parties to electoral defeat through their selection of candidates. Using comprehensive hand-collected biographical data and machine-learning algorithms to classify over 2000 candidates, I find evidence of the Conservatives choosing candidates from a wider range of backgrounds in response to electoral defeat, and evidence of the targeted use of candidates to win votes. These results may have implications for labour market inequalities in other occupations, as well as for policy decisions.

Keywords: Political Representation, Machine-Learning Classification, Inequality, Voting and Elections

JEL Codes: D72, D83, J23, J70, N43

[†]*Address*: Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD. *Email:* sml56@cam.ac.uk

'The demand of the democracy is really for a free choice of doctors. Instead of confining, as it were, those who were able to remedy their evils and cure them to a small class, they say "We want an unlimited choice in picking out those who suit us best".'

David Lloyd George, 10th August 1911[1]

1 Introduction

Although the right to vote is near-universal in many countries, and has been so for many years, there still remain significant inequalities in political representation. These inequalities exist over gender, racial, and socioeconomic lines, among others. For example, only two percent of US Members of Congress are from working-class backgrounds, as well as only five percent of MPs in the UK House of Commons, despite making up around half of the population in both countries (Carnes 2012). However, there have also been recent signs of change. For example, parties may select candidates from a wider range of backgrounds to appeal to a wider set of voters, like the Republican Party in the recent 2020 US congressional elections[2].

Understanding the causes of such inequality in political representations may have implications beyond politics, such as for inequalities in the labour market. A party's decision to select candidates mirrors the decisions made by firms hiring workers and organisations admitting members. Firms may choose to be more inclusive in their hiring for similar reasons that political parties decide to select candidates from a wider range of backgrounds. I suggest that greater inclusivity may be driven by organisations learning from failure.

Further, such inequalities in political representation may have a profound impact on the policies that are chosen. There is extensive empirical evidence suggesting that the characteristics of politicians influences their decisions, whether it be their race (Hopkins & McCabe 2012, Beach et al. 2019), gender (Clots-Figueras 2011, Bhalotra & Clots-Figueras 2014), and socioeconomic class (Carnes 2012), or even whether they smoke (Burden 2007), serve in the military (Gelpi & Feaver 2002), or have daughters (Washington 2008). As a result, inequality of political representation may have a significant effect on policy choices and subsequently socioeconomic outcomes, such as inequality.

One important contributing factor to inequalities in political representation may be the deci-

---

[1]Hansard 10th August 1911 Volume 29 Columns 1365-1477 'Debate on the Payment of Members'

[2]As reported in the New York Times: 'Republican women delivered critical victories to their party in the election, signalling the success of their efforts to recruit and elect a more diverse slate of candidates' on 4th November 2020: www.nytimes.com/2020/11/04/us/politics/republican-women-congress.html

sions made by political parties in their selection of candidates for election. Political parties often have considerable campaigning resources and so have influence over the selection of realistic candidates for election, effectively gatekeepers to public office.

The main contribution of this paper is to suggest that political parties may be strategic about the types of candidates that they select, and how this influences political representation, of which there is little theoretical or empirical discussion in the economics literature. I investigate this by identifying whether parties learn from a poor performance and choose candidates from a wider range of backgrounds to win the next election. I choose to focus on the Conservative Party's response to the shock landslide defeat in the UK 1906 General Election. While examples of parties using candidate profiles to win elections are commonplace, this setting uniquely allows for the identification of the influence of electoral incentives on a party's choice of candidates.

I also exploit a comprehensive range of candidate characteristics, which I hand-collect from not only the biographies of MPs but also unsuccessful candidates. These characteristics include information about the backgrounds of these candidates including their education, their family's backgrounds, their previous occupations, their social connections, clubs and hobbies, and so on.

In addition, I implement machine learning algorithms (Support Vector Machine and Latent Dirichlet Allocation) in a novel application: to sort candidates into different socioeconomic groups. This reduces the prior judgement required and ambiguity in the identification of candidates as political elites (for example those families have a long history of involvement in politics) or outsiders (such as the working class), or somewhere in between.

I use geographic variation in the Conservative vote share and control for the party's past performance and past candidates. I find evidence that the greater the defeat, the more likely the Conservatives were to switch from candidates from the political elite to those from working-class backgrounds. In the average constituency, a one standard deviation increase in the Conservative defeat margin in 1906 is associated with a decrease of 7.9-10.0 percent in the likelihood that the candidate in 1910 is from the political elite. The same increase is associated with a 19.3 percent increase in the likelihood of the Conservative candidate being from the working class.

I find that this switching is unique to the Conservative Party, consistent with historical evidence of the Conservatives learning slower about voter preferences than other parties. In fact, this effect is largely driven by the replacement of candidates who lost in 1906 rather than

3

replacing incumbent MPs in the subsequent 1910 elections. In addition, the Conservatives switched away from the political elite more aggressively in constituencies where they had traditionally performed well, where a given margin of defeat would have been perceived as a greater shock than in other constituencies. Further, this switching is more prevalent and aggressive in working-class constituencies. This is consistent with parties using candidates as a targeted instrument to win votes, as explained in the conceptual framework. Finally, I find that this targeted switching was beneficial to the Conservatives in the subsequent elections in 1910.

Contributions to the Literature I primarily contribute to the literature analysing political under-representation of different groups in society, and especially the causes of such under-representation. There is relatively little work exploring political representation across different socioeconomic groups, such as the working class. Carnes (2012) finds that working-class legislators in the US House of Representatives tend to mirror the economic preferences of working-class voters. In addition, their exclusion from legislatures fuels distrust of political institutions (Barnes & Saxton 2019) and decreased political engagement (Heath 2018). As for why there are relatively few politicians from poorer socioeconomic backgrounds, Carnes (2016) argues that the limiting factor is not lack of qualifications or any other supply-side issue, but political parties' low demand for candidates from a wider range of backgrounds. Further, initiatives to increase supply, such as increases in politicians' salaries, in fact crowd out the working class as it increases the supply of middle and upper class candidates (Carnes & Hansen 2016). I also contribute to a wider literature concerning the representation of historically under-represented groups in politics, such as women and ethnic minorities. There is a considerable body of work that investigates both the causes of such under-representation, such as the bias of voters (Bhalotra et al. 2018, Fulton 2012, Kahn 1996) and political parties (Sanbonmatsu & Dolan 2009, Anastasopoulos 2016, Folke & Rickne 2016), as well as the consequences of under-representation of women (Clots-Figueras 2011, 2012, Bhalotra et al. 2018, Catalano Weeks 2019) and ethnic minorities (Hopkins & McCabe 2012, Beach et al. 2019) in politics.

I add to the literature by proposing and finding empirical evidence for a different factor impacting political representation of traditionally under-represented groups: electoral pressures, where parties feel they need to select candidates from a wider range of backgrounds to win an election. In addition, the relatively fixed supply of candidates from non-elite backgrounds allows for the identification of parties' demand for wider representation among its candidates. Focusing on the demand-side decision of parties selecting candidates disentangles the supply and demand for political representation that the literature generally struggles to pull apart.

4

Further, I use machine-learning algorithms on comprehensive bibliographical data. As opposed to the literature which focuses on one characteristic of a candidate (for example gender or occupation), I use a wide array of characteristics to indicate whether a selected candidate part of the political elite or not.

Secondly, I contribute to the literature that discusses the link between political participation and policy, and more specifically whether wider political participation leads to more egalitarian policy. Wider political participation may mean extending the suffrage as well as increasing political participation among those already eligible to vote. Downs (1957) and Meltzer & Richard (1981) argue that the composition of the electorate may have direct implications for policy, the latter suggesting that if the franchise is extended to those with lower incomes there may be greater appetite for redistribution. Various work has suggested that the extension of the franchise has a positive influence on redistribution (Kotera & Okada 2017, Aidt et al. 2006, Aidt & Jensen 2013, Boix 2003), confirming the logic of Meltzer & Richard (1981). Indeed there is more recent evidence of how political participation of other groups may influence policy: Jensen & Yntiso (2019) find that the Jim Crow laws leads to a larger decrease in social spending in counties with a greater African-American population, while Bertocchi et al. (2020) find an increase in youth participation at elections, through preregistration laws, is associated with greater public spending on education. However, Aidt et al. (2020) exploit structural breaks in fiscal policy and suffrage in the United Kingdom and find little evidence of a link between franchise extension and fiscal expansion.

I contribute to this literature by providing a stepping stone between wider political participation and policy changes: political representation. I argue that even if wider political participation may increase *demand* for, say, more egalitarian policies, if it is not matched by an increase in *supply* of those policies by political parties there may be little change in policy. Given the significance of politicians' characteristics on their decision-making, as is the consensus in the political economy literature, wider political participation may only have a meaningful impact on policy once there is wider representation among politicians. As I find, elections may accelerate this process by putting pressure on parties to choose candidates from a wider range of backgrounds.

Therefore this is also a contribution to the literature concerning political competition, for example Downs (1957) and Lindbeck & Weibull (1987). While these papers argue that political parties compete using policy platforms, I suggest that they may also compete with the portfolio of candidates that they select.

Paper Layout The paper proceeds as follows: Section 2 explains the conceptual framework from which the empirical strategy is motivated; Section 3 outlines the relevant institutional context surrounding the 1906 election and the Conservative response. Section 4 describes the comprehensive electoral and biographical information collected; Section 6 explains the baseline empirical strategy as well as the machine learning algorithms that sort the candidates into socioeconomic groups (Support Vector Machine and Latent Dirichlet Allocation); Section 7 reports the baseline results. Section 8 reports further evidence of learning-by-losing; Section 9 shows the effect of this strategy on outcomes in the 1910 elections; Section 10 concludes.

2 Learning-By-Losing: Using Candidates to Win Elections

In this section, I provide a conceptual framework motivating why political parties may use their portfolio of candidates, and their characteristics, as a targeted instrument for winning elections. I then argue that political parties may change the portfolio of candidates as a result of losing an election, which may widen representation among their candidates. This provides motivation for the empirical strategy outlined in the following sections.

2.1 Why do voters care about candidates?

A candidate's profile or characteristics may influence electoral outcomes for two reasons. The first is that voters may feel more positively about a candidate who shares some characteristics with them. This may even be independent of the policy platform of the party that they represent. This is consistent with social identity theory from the political science literature (Tajfel 1974, Ashforth & Mael 1989, Conover 1984, Shayo 2009) as well as recent empirical work, for example Ansolabehere & Puy (2016). The concept of identity has also been discussed in the economics literature (Akerlof & Kranton 2000, Besley & Persson 2019), where voters may choose a candidate or party even if it is not in their economic interests to do so.

Secondly, voters may observe the characteristics of a candidate as a commitment to future policy on behalf of the party that chose them. For example, if a party chooses a candidate from the working class to run for office, it may be a signal to voters that the party intends to propose policies that are more favourable to the working class in the future. There is empirical evidence suggesting that the characteristics of a politician may influence their policy preferences (Clots-Figueras 2011, Bhalotra & Clots-Figueras 2014, Hopkins & McCabe 2012). If a voter observes a candidate with similar characteristics to them, they may interpret this as a sign that the party they represent are more likely to choose policies that benefit them in the future.

This may be the case irrespective of current policy platform. A party may not be able to credibly change policy platform radically, so a change in candidate profiles may be a commitment device for a different policy platform. For instance, a party may choose a working-class candidate to show its commitment to working-class causes in the future, as changing the manifesto to include more economic redistribution may not be credible as such promises may be undone overnight[3]. This is amplified in the case of a significant incumbency advantage (Lee 2008, Hainmueller et al. 2015), where choosing a candidate opens the possibility of that candidate being re-elected in the future, so is a credible long-term commitment.

2.2 Using candidates to win elections

Parties may take advantage of voter preferences by adjusting their portfolio of candidates across constituencies. Further, this may be a targeted instrument that can be adjusted to win votes from different types of voters. The political economics literature focuses on parties changing their platforms to win elections (Downs 1957, Lindbeck & Weibull 1987), I argue that the portfolio of candidates is another part of their strategy.

Further, changing the portfolio of candidates may be a more targeted instrument of winning votes across constituencies than the policy platforms. A policy platform may be a national manifesto that is delivered across the country. In this case, a change in the policy platform may lose some votes as well as win others. However, there may be an opportunity for parties to change individual candidates and attenuate the negative spillovers of placating one set of voters.

This provides the motivation for political parties to use the portfolio of candidates to maximise their vote shares across many constituencies. Even if the policy platform can be somewhat tweaked geographically, there will be an incentive to use candidates as a more targeted instrument so long that the negative spillovers of changing a candidate in one area on votes won in other areas relatively low. This seems reasonable as voters are likely to care more about the candidates running in their own constituency than candidates running in other constituencies, even those nearby.

2.3 Learning-By-Losing: Empirical motivation

While political parties may choose their portfolio of candidates to maximise electoral success across many constituencies, they may also do this by a learning process than spans across

---

[3]This is similar to the logic of Acemoglu & Robinson (2000).

elections[4]. Even if they have complete information about the characteristics of their candidates, they may have imperfect information about voters' preferences. For example, they may not have perfect information about voters' preferences for a certain types of candidates.

As a result, political parties may undergo a learning process that may be characterised as 'learning-by-losing'. A party chooses a portfolio of candidates before an election. If the party performs well at the election, they may choose a similar portfolio of candidates for the next election. However, if they perform poorly, they may make changes.

The economics literature concerning experimentation by individuals (Keller et al. 2005) and groups (Strulovici 2010, Gieczewski & Kosterina 2019, Gieczewski 2020) suggests that they may tend to 'under-experiment', whether due to risk aversion or free-riding from others' experimentation. If political parties also exhibit this consistent under-experimentation, it may take especially poor election results for them to finally experiment.

A political party may therefore change the type of candidate in a constituency in response to a poor result. The greater the defeat (compared to what was expected), the greater the signal to the party that they should change their strategy. If this is the case, one would expect that a poorer performance in a constituency in one election should make changing the type of candidate at the next election more likely. One may expect that a result of this learning-by-losing process could result in the selection of candidates from a wider range of backgrounds, especially if the party were initially choosing from a relatively narrow selection of characteristics.

3 Political Representation and the Landslide 1906 Election

In this section I outline the relevant political and historical background of British politics between 1900 and 1910. I briefly describe the political institutions, as well as the historical context of the four elections between 1900 and 1910. I explain why the shock landslide defeat of the Conservative Party in 1906 provides an ideal setting to isolate the decision of the party to select candidates from a wider range of backgrounds in response to the electoral defeat.

3.1 Landslide Conservative defeat in 1906

By the beginning of the twentieth century, British politics had been dominated by two political parties (the Conservatives and the Liberals) for decades. Partly helped by the early popularity

---

[4]They may also change their policy across time, but I focus on changes to the profiles of candidates. Further, I argue in the previous sub-section why political parties may prefer to change the portfolio of candidates rather than the policy platform.

of the Boer War, the Conservatives won a large majority of 130 seats (in a House of Commons of 670 members[5]) at the 1900 election. However, the 1906 election saw the Conservatives lose 246 seats in a landslide defeat that left even the Prime Minister Arthur Balfour without a seat in Parliament. As Blewett (1972) suggests, this landslide defeat was the result of a perfect storm of frustration in the Boer War and more general concerns over poverty across the United Kingdom, emphasised by an embryonic trade union and Labour Party movement. Further, the Rowntree Report in 1899 brought the scale of poverty among the working class to the attention of the political elite and the middle class, as well as reports that forty per cent of recruits for the Boer War were so malnourished and unhealthy that they were deemed physically unfit for war. Sykes (1998) suggests that the Conservatives viewed the primary factor behind the landslide defeat in 1906 was 'the desertion of the Conservative working-class voter.' The next few years would see unprecedented redistributive policy under a Liberal government, which saw the introduction of a welfare state system[6]. In contrast, the Conservatives opposed these policies, as they had done in 1906, emphasising the trade-off between a welfare state and Britain's ability to fund its army and navy. The Conservative-dominated House of Lords rejected the 'People's Budget' of 1909, and the Liberal government were forced to hold not one but two elections in 1910: the first in January to show support for the Budget, while the second in December was later called to show support for removing the House of Lords' ability to reject budgets.

3.2 Elections in Britain 1900-1910

The House of Commons throughout the four relevant elections (1900, 1906, January 1910, and December 1910) consisted of 670 members. The vast majority of Members of Parliament (MPs) were the sole representative of their constituency, with a minority of constituencies represented by two members. In single-member constituencies, political parties would select one candidate for the general election. As Blewett (1972) notes, the central party organisation had considerable influence over the selection of candidates, in particular through campaign funding.

---

[5]The United Kingdom Parliament is made up of two legislative chambers: the House of Commons which is made up of elected officials, and the House of Lords which is made up of appointed and hereditary peers.

[6]The Liberals introduced a raft of unprecedented welfare policies: the Education Act (1906) providing new free school meals; the first nationwide state pensions and unemployment insurance were introduced in the Old Age Pension Act (1908) and the National Insurance Acts (1911); the Trade Boards Act (1909) set minimum wages in some trades; job centres were set up in the Labour Exchanges Act (1909); notably the Trade Disputes Act (1906) was the first significant legislation that provided labour unions some protection from their employers when striking.

3.3 Why focus on the 1906 election?

The shock landslide election in 1906 provides a unique opportunity to focus on the choice of the party in selecting candidates from a wider range of backgrounds: firstly the unexpected nature of the 1906 election result, and secondly the Conservatives' highly centralised process of selecting candidates. These suggest that changes to the type of candidates run in 1910 would be the result of the Conservatives learning from the 1906 election and selecting from a wider range of backgrounds to do better in the 1910 elections.

3.3.1 Unexpected shock of the 1906 landslide

The first advantage of this setting is the unexpected result of the 1906 election. The Conservatives lost 246 out of their 402 seats (in a 670 seat House of Commons), after eleven years of dominance. The popularity of the Liberal Party's policy of radical redistribution was unexpected at the 1906 election, despite the Boer War and the Rowntree Report highlighting the extent of poverty within the United Kingdom[7]. However, this was also a landslide that was indiscriminate across regions and socioeconomic classes, in constituencies where they had previously been popular as well as those where they were unpopular[8].

Figure 1: Conservative 1906 and 1900 Election Results across Constituencies

---
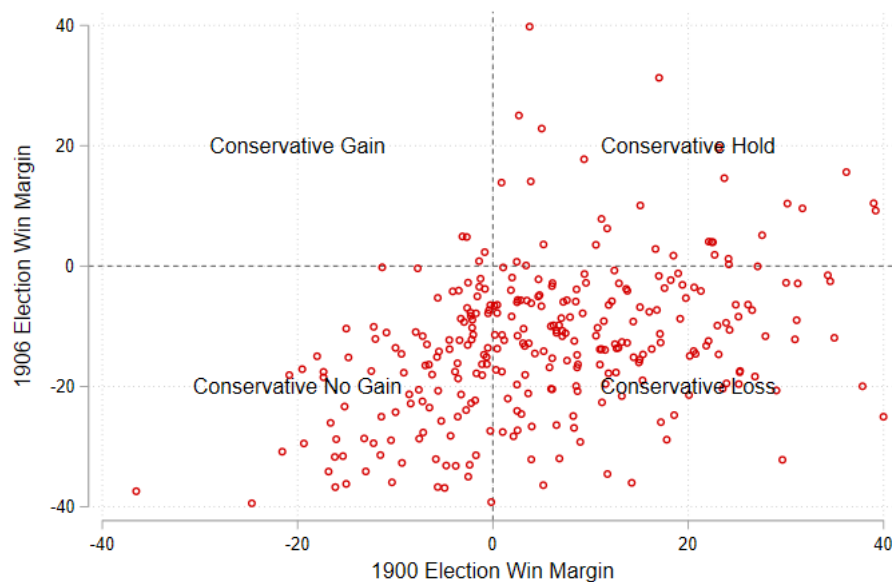
[7]The Rowntree report in 1899 brought the scale of poverty among the working class to the attention of the middle class, while the Boer War (1899-1902) shone a disturbing light on the health (or lack of) of the country: 40 per cent of recruits were determined unfit for duty. Although there were many other reasons for the Conservative defeat in 1906. Blewett (1972) provides a more detailed explanation of the political timeline.

[8]This is illustrated by the fact that the swing against the Conservatives in 1906 was 12.2 per cent of the vote among predominantly upper class constituencies; 11.7 per cent among middle class constituencies; 13.1 per cent for working class constituencies.

This is presented in Figure 1, which reports the Conservative win margin (as a proportion of total votes) in 1906 compared to that of 1900. This election result was also a shock. All but three of the Conservative cabinet lost their parliamentary seats, including Prime Minister Balfour himself. Had the margin of defeat been known, it seems likely that more resources would have been thrown into assuring the re-election of significant Conservative Party members. Unlike in the previous century, there had been no recent extension of the franchise to new voters. Further, the swing away from the Conservatives in 1906 (5.4 per cent of total votes) was unprecedented in British politics.

Opinion polling before the election was virtually non-existent in the early twentieth century, with politicians relying on often idiosyncratic results of by-elections instead, limiting the ability of the Conservatives to anticipate the result. This is echoed by the response to the result, with The Times reporting that the Liberals *'will realise all that they dreamt and hoped for, but scarcely ventured to expect ... about to suffer an embarrassment of riches.'*[9] This election result, in its magnitude and unexpectedness, provide an ideal quasi-exogenous shock to the Conservative's selection of candidates in the subsequent elections.

3.3.2 Strong party control over candidate selection

The second advantage of this setting is the relatively fixed supply of candidates from non-elite backgrounds, with selection controlled by the central party. Before the Parliament Act of 1911, Members of Parliament (MPs) are not paid and therefore required to pay their own expenses, as well as a large proportion of their campaign funding (Blewett 1972). Even if more candidates from non-elite backgrounds wanted to run for office in light of the 1906 election, they would not be able to do so unilaterally: the supply of candidates from non-elite backgrounds is relatively fixed. The average candidate in a contested seat spent £1,100 in campaign financing (Blewett 1972), while the average annual income in the United Kingdom in 1900 was £46.64 (Allen 1994). Candidates from non-elite backgrounds therefore rely on central party funding decisions, such as the £13,940 spent by the Conservative Party on working-class candidates in the 1910 elections.

If the result of the 1906 election prompted the selection of candidates from a wider range of backgrounds, it is likely the consequence of increased *demand* for non-elite candidates from political parties rather than *supply* through non-elite candidates bankrolling themselves.

Overall, the 1906 election provides an opportunity to interpret a change in candidate selection

---

[9] The Times Newspaper, 16th January 1906.

as a result of the Conservatives learning from losing the 1906 election. This setting provides a unique opportunity to identify political party's decisions in selecting candidates to try and win elections.

4 Descriptive Summary of Electoral and Biographical Data

4.1 Electoral and Constituency Data

The unit of observation is a party-constituency-election (for example, the Conservative Party candidate for the City of London constituency in the 1906 General Election). The 3,938 party-constituency-election combinations cover four UK general elections (1900, 1906, January 1910, and December 1910) in constituencies in England, Scotland, and Wales.

The electoral data is extracted from Craig (1974) *British Parliamentary Election Results: 1885-1918*. This provides the names of candidates (in some cases their title), their party, the number of votes for each candidate, and the rate of turnout. Summary statistics about the elections can be found in the Appendix A, but the most important aspect of the electoral data is documenting the extent of the Conservative loses in 1906 (compared to 1900). In addition, information about constituency characteristics is taken from Blewett (1972), which is presented in Table 1.

Table 1: Constituency Characteristics

Population	68,377
Number of Houses	13,047
Urban	48.8%
Predominantly Upper/Middle Class	13.3%
Mixed Classes	24.6%
Predominantly Working Class	22.6%
Mixed Urban/Rural	12.0%
Mining Community	2.3%

Notes: This table reports summary statistics concerning constituency characteristics, either as the mean value across constituencies (population and number of houses) or as a percentage of constituencies that match the description according to Blewett (1972).

4.2 Biographical Data

I hand-collect biographical information concerning not only concerning those who won election to the House of Commons between 1900 and 1910, but also the vast majority of unsuccessful candidates. This provides a wider insight into whether the Conservative party opened its doors to those outside of the establishment than only focusing on those who were elected.

This data is compiled primarily from two biographical dictionaries. The first is *Who's Who of British Members of Parliament* (Stenton & Lees 1979), which provides summarising biographies

of Members of Parliament from 1832 to 1945. This provides information not only about those who sat in Parliament between 1900 and 1910 (which is the period of interest in this paper), but also those who ran unsuccessfully in that period and succeeded before 1900 or after 1910. The second is the *Oxford Dictionary of National Biography*, which covers many unsuccessful candidates who played a significant role in society in other ways. Combined, these sources contribute information on 2,245 candidates across the four elections between 1900 and 1910.

The full extent of the biographical data can be explored in the Appendix A.1, but the main characteristics of each candidate are their titles (honorary, military, hereditary); their education (school and university); their previous occupations before running for office; the occupations of their family members (often their father and father-in-law's). In addition, there is data on their social connections, which illustrates to what extent the candidate was well connected to the political or social establishment. This information includes all of the social clubs they were members of; what law houses they practised at; the number of connections with other candidates in the sample they had by virtue of their school, university, law houses, and social clubs; whether they were appointed Justice of the Peace[10]. Using this wealth of information I am able to assess whether the Conservative party selected candidates in 1910 who were very different from those in 1906, as well as identify those who are from the political elite/establishment.

There are inevitably some limitations associated with the data. The most important is the lack of data for 10.7 and 11.5 per cent of all and Conservative candidates, respectively[11]. I discuss the potential impact of missing biographical data in Appendix F.4. I report results for a wide range of assumptions about those with missing biographical data, and still find similar results. Another issue is the incidence of two-member constituencies, such as Plymouth and Cambridge University, as it is unclear what the margin of loss is for a candidate or party, but this is only the case for 23 constituencies. Furthermore, biographical data is scarce for the Irish MPs (and more so for candidates in general), so their 103 constituencies are removed.

5 Machine-Learning Classification of Candidates

In this section I provide motivation for the novel use of machine-learning to classify candidates. I outline the procedures of two methods the Support Vector Machine (SVM) and the Latent Dirchlet Allocation (LDA). These provide estimated probabilities that a candidate is a member of a particular socioeconomic group (for example the political elite), fully exploiting

---

[10]This title implied they had many connections in the local area they were appointed a JP for.

[11]10.7 per cent of all party-constituency-election trios and 11.5 per cent of Conservative constituency-election pairs.

the wealth of hand-collected biographical data.

5.1 Machine-learning and candidate classification

While there is a wide range of biographical data available for each of the candidates, I aggregate this information using two different machine learning methods: a Support Vector Machine (SVM) and Latent Dirichlet Allocation (LDA).

5.1.1 Support Vector Machine (SVM)

The first method I use is the Support Vector Machine, which has been used in various scientific fields, such as bioinformatics (Manavalan & Lee 2017) and brain image classification (Zhang et al. 2015). In economics, Gründler & Krieger (2016) use a support vector machine to compile a democracy index using a range of variables, while Ghoddusi et al. (2019) discusses its popularity in predicting energy prices and consumption.

An intuitive explanation (a technical discussion can be found in Cortes & Vapnik (1995) and Gründler & Krieger (2016)) of the support vector machine process is as follows:

1. Label candidates who are unambiguously part of the political 'elite' or unambiguously 'outsiders'.

2. Assign Elite (SVM)=100 to those labelled as 'elite' and Elite (SVM)=0 to those labelled as 'outsiders'.

3. Use a support vector machine to find an optimal hyperplane that maximises the distance (within the space of all characteristics besides those used to label the unambiguous cases) between itself, the 'elite', and the 'outsiders'.

4. Use this hyperplane to categorise unlabelled candidates (continuously from 0 to 100).

For the initial labelling, I define those who are unambiguously 'elite' as candidate who have both a hereditary title in their family and a member of their family who is involved in politics, which labels 6.7 per cent of the observations. This is likely to well identify the pinnacle of the political elite, as these candidates are likely to have connections to both the House of Commons and the House of Lords. I define those who are unambiguously 'outsiders' as those whose families were traditionally manual workers/labourers, which labels 9.0 per cent of the observations. These families are highly unlikely to have a previous connection with Westminster politics - especially as the trade union movement had only just begun to involve themselves in Westminster politics at the turn of the century. While initial labelling is inherently

somewhat arbitrary, I believe this labelling is justified by the historical context. Furthermore, in Appendix F.1 I alter the labelling assumptions about the unambiguously elite and outsiders which show the robustness of the following results. The following results use the continuous form of the estimated **Elite (SVM)** variable.

5.1.2 Latent Dirichlet Allocation (LDA)

To support this analysis, I also use Latent Dirichlet Allocation. This unsupervised machine learning topic model clusters text into a number of topics, which has been used to analyse text in a variety of scientific settings, such as grouping medical studies (Wu et al. 2012). Latent Dirichlet Allocation has also been recently used in economics to analyse the transcripts of the Federal Reserve's Open Market Committee to analyse transparency in policy-making decisions (Hansen et al. 2018), CEO behaviour (Bandiera et al. 2020), and political ideologies among voters Draca & Schwarz (2019). While a thorough explanation of the method can be found in Blei et al. (2003) (and its implementation by Bandiera et al. (2020) who explain its advantages over other methods such as k-means clustering - in Appendix F.2 I use k-means clustering and find similar results), the following is a more intuitive explanation.

The LDA finds clusters of candidates across the space of characteristics, where the number of clusters is specified. The LDA then estimates the probability that a candidate is a member of one of the groups represented by the clusters. For example, one cluster might be one where those whose families have hereditary titles may also be more likely to go to Eton and then Oxford or Cambridge. Although the number of clusters are specified, the groups returned and their characteristics are not. I assume three groups[12], but in the Appendix F.3 I allow for different number of groups and find similar results[13].

The LDA is agnostic about the nature of the groups it presents, so I present three word clouds that provide some intuition. Figure 2 presents word clouds from each of the three groups where the size of the word is weighted according to the number of members[14] who have that characteristic. Candidates who are likely to be members of Group 1 have families with hereditary titles, went to fee-paying schools and went to Oxford or Cambridge for university - potentially the political elite. Likely members of Group 2 also are likely to go to Oxford and Cambridge, but have occupations more focused on the professions, such as law, the civil ser-

---

[12]The social class hierarchy of Britain in the early twentieth century also provides motivation for initially choosing three groups with the notion of an upper class, middle class, and working class (Blewett 1972).

[13]The driving force behind this robustness is the distinctive clustering of those who have hereditary titles into one group and those whose families are labourers into another, even for larger group sizes.

[14]Members being defined as candidates whose probability of being in that particular group is greater than any of the other probabilities of being in one of the other groups.

Figure 2: Word Cloud of Common Characteristics

Group 1

Group 2

Group 3

This figure presents a word cloud of common characteristics within the three groups

vice, and academia - potentially the middle class. Finally, the likely members of Group 3 are overwhelmingly those who were, as well as their family, skilled labourers - likely the working class.

As a result, I interpret the probability of a candidate being a member of group 1 as being the probability that they are a member of the political elite, and likewise for group 2 (middle class) and group 3 (working class).

5.1.3 Elite Index Variables

The two algorithms estimate probabilities between 0 and 100 of a candidate being a member of a group:

1. **Elite (SVM)**: Probability that the candidate is in the political elite, according to the SVM.

2. **Elite (LDA)**: Probability that the candidate is in the political elite, according to the LDA.

3. **Middle (LDA)**: Probability that the candidate is in the middle class, according to the LDA.

4. **Working (LDA)**: Probability that the candidate is in the working class, according to the LDA.

Elite (SVM) provides an indication of switching between candidates from the political elite to those further away from the political elite, whether it be the middle class or the working class. The variables estimated by the LDA may illustrate more nuanced patterns of switching between the political elite to the middle class, the middle class to the working class, and so on. I use all four of the estimated probabilities as dependent variables in the main analysis.

5.2 Advantages of the LDA/SVM over other classification methods

The main advantage of using these machine learning algorithms is that it reduces prior judgement required in the classification of candidates. Manual classification of candidates would either be somewhat arbitrary or make assumptions about whether a certain characteristic belongs to one group or another. While some candidates are fairly obviously part of the political elite or the working class, there may be many cases where it is less obvious.

Further, the extent to which a certain characteristic makes a candidate more or less likely to be part of a group may not be obvious. Whereas having family who were involved in politics and have hereditary title may be an obvious characteristic of those in the political elite, whether a candidate was involved in local politics may not be as obvious an indicator[15]. As argued by Gründler & Krieger (2016), these machine-learning methods with endogenous weighting on different characteristics avoid the problem of arbitrary weighting that other aggregation methods may suffer from, while still making use of the full extent of the biographical data.

The SVM mostly only relies on fairly intuitive labelling of some observations. Further, the

---

[15]For example, the political elite may be involved in local politics because they are influential in their local area, whereas local politics may also include those from working-class backgrounds who were elected to local councils.

LDA only requires the specification of the number of groups, which I vary in the Appendix, especially as I include all characteristics. Draca & Schwarz (2019) outline some advantages of using the LDA compared to other classification methods: firstly that it allows for non-linear relationships between characteristics compared to Principal Component Analysis (PCA) and Factor Analysis (FA) which do not. For example, the LDA would allow for the fact that, given other characteristics, being involved in local politics may be an indicator of being part of the political elite (because they were influential in their local area already) or the working class (because for them it was their primary avenue into politics). Further, compared to k-means clustering and Latent Class Analysis, Draca & Schwarz (2019) argue that the LDA allows for mixed membership - for example 10 percent political elite, 40 percent middle class, 50 per cent working class - rather than just a single probability that a candidate was a member of one group.

6   Empirical Strategy

I use the following baseline empirical strategy to find whether the Conservatives responded to the electoral shock of 1906 by changing the identity of the candidate for the 1910 elections:

$$y_{i,1910} = \alpha + \beta_1 LossMargin_{i,1906} + \beta_2 y_{i,1906} + \beta_3 LossMargin_{i,1900} + \epsilon_{i,1910}, \tag{1}$$

where $y_{i,1906}$ and $y_{i,1910}$ are values of the outcome variable, which is the probability that the Conservative candidate in constituency i is a member of a certain group which is a certain characteristic, in the 1906 and the 1910 elections, respectively; $LossMargin_{i,1906}$ and $LossMargin_{i,1900}$ are the margins by which the Conservative party lost the election in constituency i in the 1906 and the 1900 elections, respectively. I define the variable by the loss margin so that the explanations of the results are more intuitive: for avoidance of doubt I also include observations with negative values of $LossMargin$ (where the Conservative won).

I include $y_{i,1906}$ so that β_1 can be interpreted as the effect of the 1906 election on candidate characteristics in 1910 *relative* to the candidate in 1906. The coefficient of interest is β_1 - this estimates how much the Conservative party candidate changed for the 1910 elections in response to electoral failures in 1906. I also include $LossMargin_{i,1900}$, so that the variation across constituencies is how poorly the Conservatives did *relative* to their past performance in 1900. One would expect that, if the Conservatives lost by the same margin in two constituencies, the shock would be greater in the constituency where historically they had performed well. This

is robust to other controls for previous Conservative performance, such as how many elections the Conservatives won in the constituency between 1885 and 1900. The treatment group is effectively constituencies where the Conservative lost by a large margin in 1906 relative to their previous performance in 1900, the control is a loss margin in 1906 that is not as large relative to 1900, compared to other constituencies.

Another reason for including $y_{i,1906}$ and $LossMargin_{i,1900}$ is to control for unobserved constituency heterogeneity. While the 1906 election was a shock to the political elite, the geographical variation in this electoral shock, the variation in the Conservative loss margin, may not be exogenous (although Figure 1 suggests that heavy losses were experienced across seats with a wide range of 1900 performances). For example, the $LossMargin_{i,1906}$ may be low in constituencies where the Conservatives have previously performed well, which may be associated with long-incumbent Conservative MPs who are more likely to be part of the political elite. I include these controls in the remainder of the paper (estimating standard errors clustered at the constituency level), and Appendix E reproduces these results without controls which lead to similar conclusions.

7   Results

In this section I report the baseline results, which provides evidence of the Conservatives responding to the 1906 election by switching representation among their candidates from those in the political elite to those from the working class. I undertake a falsification test which shows that the effect is unique to the Conservatives.

7.1   Baseline Results

Table 2 reports the baseline results, and illustrates the significant effect of the 1906 defeat on the type of candidates the Conservatives fielded in 1910. Columns 1 and 2 report the effect of the 1906 loss margin on the probability of whether the Conservative candidate in 1910 is from the political elite, as estimated by the SVM and LDA, respectively. Columns 3 and 4 report the same effect on the probability of the candidate being from the middle class and the working class groups as estimated by the LDA.

Columns 1 and 2 clearly show a significant switching away from political elite. For every percentage point increase in the defeat that the Conservatives faced, there is around a 0.25 to 0.3 percentage point decrease in the probability that the candidate in the 1910 elections was from the political elite. While there is little evidence of switching to middle-class candidates

19

Table 2: Baseline Results

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.289***	-0.255***	0.068	0.193***
	(0.098)	(0.088)	(0.081)	(0.064)
1900 Loss Margin	0.046	0.009	0.069	-0.070
	(0.096)	(0.086)	(0.078)	(0.062)
Lagged Dependent Variable	0.270***	0.430***	0.351***	0.378***
	(0.044)	(0.041)	(0.047)	(0.041)
Adjusted R^2	0.089	0.205	0.108	0.162
N	468	468	468	468

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The 1900 and 1906 loss margins are the Conservative losses to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). The lagged dependent variables are the values of the dependent variables in 1906. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

in Column 3, there is significant evidence of switching to working-class candidates in Column 4. These results are also illustrated in Figure 3, which illustrate the percentage point change in the probability of the 1910 candidate being a member of a certain group, in response to a one standard deviation (17.3 percentage point) increase in the 1906 defeat.

Figure 3 shows that the one standard deviation increase in the 1906 defeat leads to a 5.9 and 5.1 percentage point decrease in the probability that the candidate is from the political elite, according to the SVM and LDA, respectively. In the average constituency, this is a decrease of 10.0 and 7.9 percent. The same increase in the defeat in a constituency is also associated with a 3.9 percentage point increase in the chances of the 1910 candidate being from the working class, which is an increase of 19.3 percent in the average constituency. The baseline results provide evidence of significant switching of candidates in response to a greater loss in the 1906 election.

I also investigate the effect of the 1906 defeat on the type of candidates the Conservatives fielded in 1910 by focusing on individual characteristics. The results of these are produced in the Appendix, illustrated in Figure D.1 and reported in Table D.1. These suggest the same conclusion: an increase in defeat in 1906 is associated with the candidate in 1910 less likely to have characteristics that are intuitively associated with the political elite (such as an Oxbridge education), and more likely to have characteristics associated with the working class (such as having a father who was a manual labourer). The following results in the remainder of the

Figure 3: Baseline Results
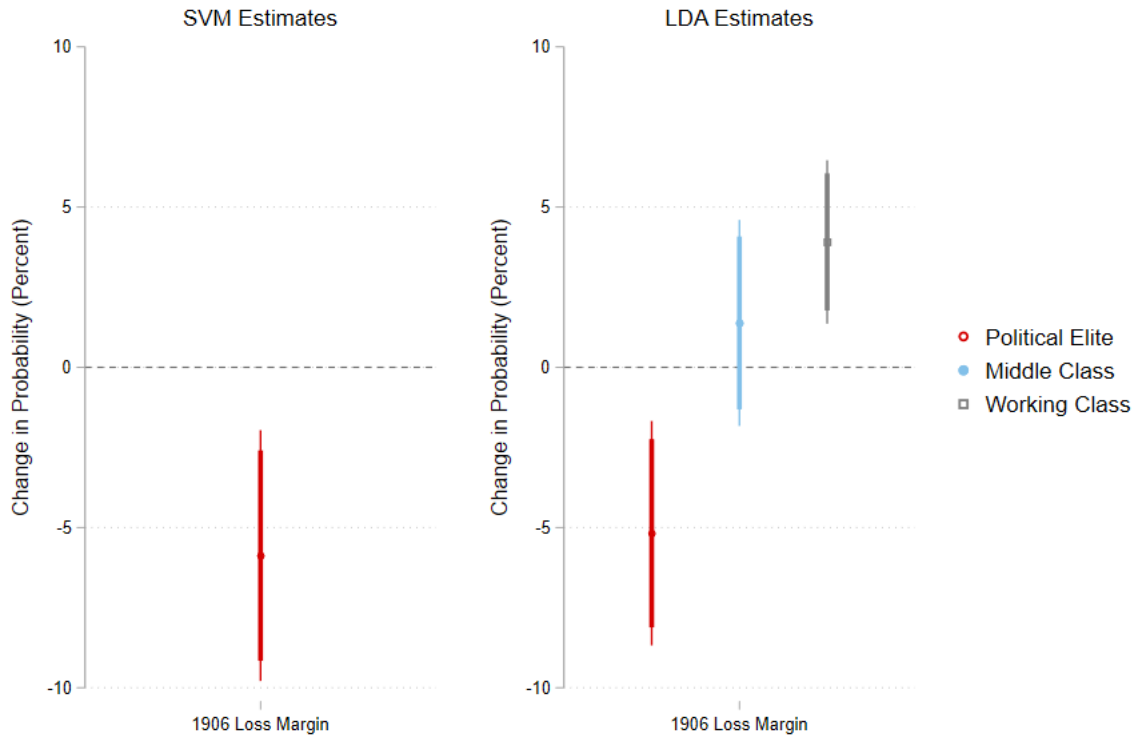
This figure illustrates the main results from Table 2: the effect of a one standard deviation increase in the 1906 loss margin on the percentage probability that the candidate is from a certain group. Both 90 and 95 percent confidence intervals are illustrated.

paper, which use the estimated SVM/LDA probabilities as the dependent variable, are also robust across individual characteristics, as reported in Appendix D.

7.2 Falsification Test - Effect Unique for Conservatives?

There is historical evidence suggesting that the other significant political party, the Liberals, had more information about voter preferences going into the 1906 election, for example the increased appetite for more redistributive policies (Blewett 1972). The Liberals had made an election pact with the Labour movement in 1903. After winning the 1906 election they proceeded to enact the introduction of the first ever statewide pensions, unemployment benefits, free school meals, trade union protection, and so on.

Therefore one would expect the 'learning-by-losing' process in 1906 and beyond to have been unique to the Conservatives: the historical evidence suggests that an outreach to more working-class voters was a lesson that had been learned already by the Liberals before 1906. As a result, the Liberals provide a falsification test for the Conservative learning-by-losing process. To un-

dertake this falsification test, I include all candidates from all parties (which also includes some Labour candidates among other smaller parties), and define the loss margin as the loss margin for the party that selected the candidate. In addition, I include an interaction between the 1906 loss margin and a dummy variable indicating whether the candidate is from the Conservative Party or not.

Figure 4: Including All Parties - Unique Effect for the Conservatives
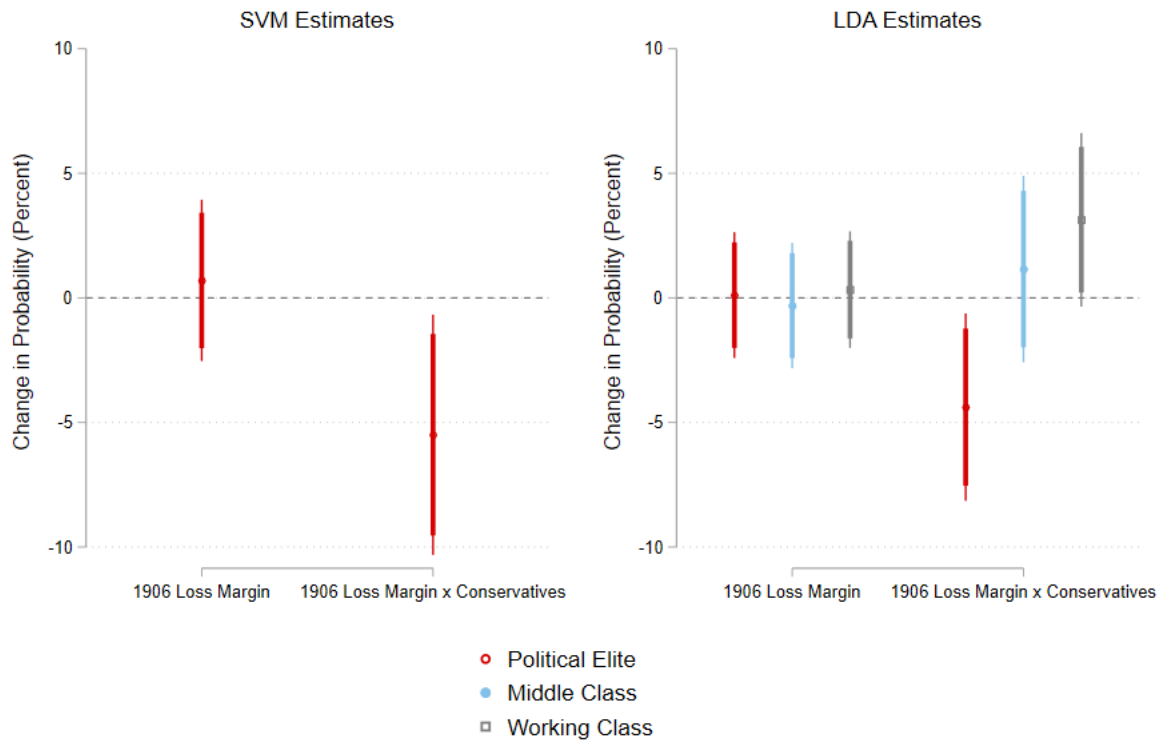
This figure illustrates the main results from Table C.1: the effect of a one standard deviation increase in the 1906 loss margin on the percentage probability that the candidate is from a certain group. Both 90 and 95 percent confidence intervals are illustrated.

The results are shown in Figure 4. The right-hand side illustrates the Conservative switching of candidates from the political elite to the working class in response to a greater loss, as in Figure 3. The left-hand side of Figure 4 confirms that this response to a greater loss margin is unique to the Conservatives. An increase in the 1906 loss margin for the Liberals (and other parties) does not appear to change the type of candidates that they select for the 1910 election. This is consistent with the learning process described above: one where the Conservatives were slower to learn about voter preferences than other parties.

8 Further Evidence of a 'Learning-By-Losing' Process

This section provides further evidence of a learning process that the Conservatives underwent. I exploit constituency heterogeneity: firstly by the discrete variation in whether the Conservative candidate in 1906 won or not. Secondly, I explore variation in the perception of the defeat. I argue that a large loss in a constituency where the Conservatives had traditionally performed well would have been a greater signal to change strategy. Thirdly, I explore socioeconomic variation across constituencies to investigate whether the Conservatives used candidates as a targeted instrument, as argued in Section 2.

8.1 Don't Lose, Don't Learn?

I further analyse the learning-by-losing process by investigating switching of candidates in constituencies where the Conservatives were defeated in 1906 and those where the Conservatives won in 1906. I argue that, in a given constituency, the signal for the Conservatives to change their candidate may be the margin of defeat in 1906, controlling for previous performance.

However, one may expect a difference in the learning-by-losing process between constituencies where the Conservatives won and lost in 1906. For example, a given change in the margin may not be perceived as the same strength of signal, depending on whether the Conservatives won or lost the election in 1906. The increase in the loss margin from 10 to 20 may be perceived as a greater signal than the same increase in the loss margin from -20 (where the candidate had previously won by 20 percent) to -10. This may be because the signal may not be perfectly continuous across the loss margin: it is easier for the Conservatives to ignore a decrease in win margin if the candidate remains in the House of Commons, and less easy for the Conservatives to ignore an increase in the loss margin if that seat is occupied by the Liberals.

Further, the Conservatives may be risk averse, and this may influence how they change their strategy in response to a signal. Even though a candidate has won by a smaller margin in 1906 compared to 1900, the party may be reluctant to uproot the incumbent as the quality and popularity of the alternative candidate is not perfectly known. Even if the signal is just as strong in constituencies where the Conservatives won, they may be reluctant to change strategy when at the very least the current candidate was elected into the House of Commons. To test this, I include an interaction between the 1906 loss margin and a dummy variable indicating whether the Conservatives lost the election in the constituency in 1906 (as well as the dummy variable alone) to the baseline specification.

Figure 5: Don't Learn Unless They Lose - 1906 Constituency Victory or Loss
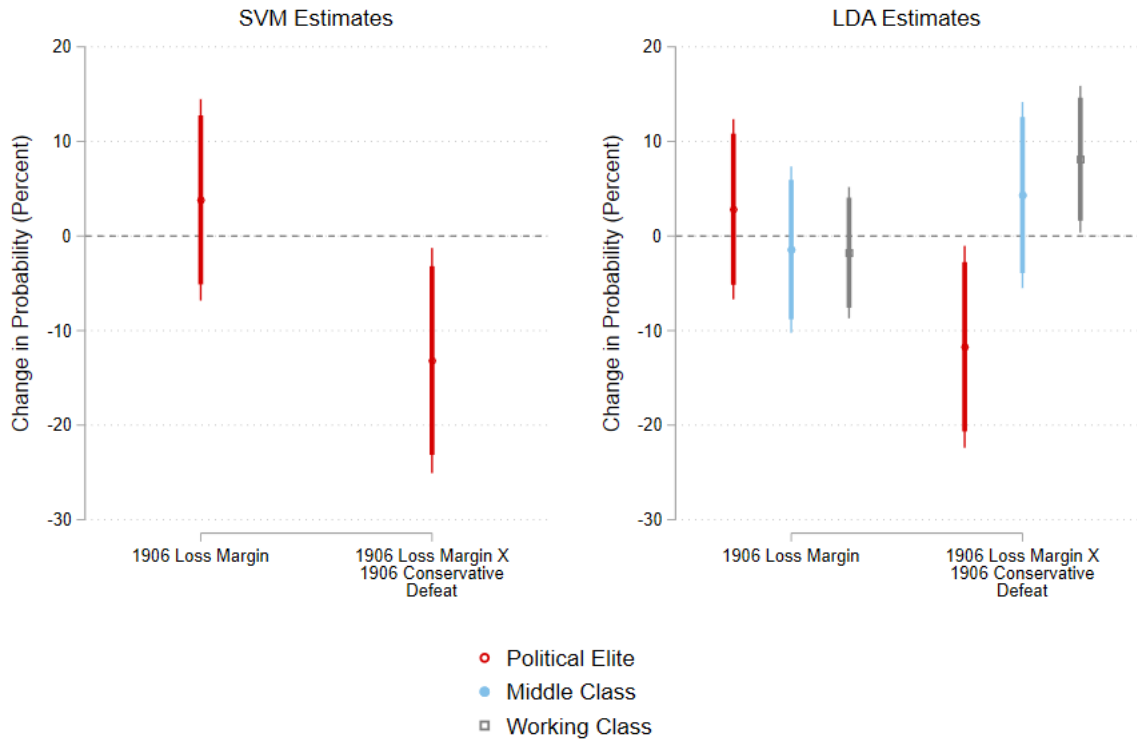
This figure illustrates the main results from Table C.2: the effect of a one standard deviation increase in the 1906 loss margin on the percentage probability that the candidate is from a certain group. Both 90 and 95 percent confidence intervals are illustrated.

Figure 5 illustrates the results, which show that the switching of candidates from the political elite to the working class (and to extent the middle class) is driven entirely by constituencies where the Conservatives lost in 1906, depicted on the right-hand side. While a smaller victory than before does not result in the switching of candidates, a greater defeat results in widening representation. In constituencies where the Conservatives lost in 1906, a one standard deviation increase in the defeat is associated with a 11.2 or 10.0 percentage point decrease in the probability that the candidate is from the political elite (19.0 and 15.5 percent decrease for the average constituency) as defined by the SVM and LDA, respectively. This also increases the probability that the candidate is from the working class by 6.9 percentage points (34 percent for the average constituency). These are significantly larger than the average effect across all constituencies presented in the baseline results. The discontinuity of the effect reduces concerns about the baseline result being driven by long-term changes in preferences across the country.

There may be other practical reasons for this discontinuity, such as the party's wish to hold onto an incumbency advantage (Lee 2008, Hainmueller et al. 2015) or to maintain party unity

which may be disrupted by replacing an incumbent politician. However, these results provide evidence consistent with a learning process where the Conservatives do not learn unless they lose.

8.2 Conservative Strongholds: A Greater Shock?

I next investigate whether the same defeat (relative to the 1900 result) is perceived as a stronger signal if the Conservatives were historically strong in the constituency. For example, a given defeat may be seen as more concerning if the Conservatives had traditionally performed well in the constituency, as it may have been a long time since the central party had to consider a new candidate for the seat. This stronger signal may be more likely to result in a change in candidate, as a product of the learning-by-losing process.

Figure 6: Conservative Strongholds and the 1906 Loss Margin
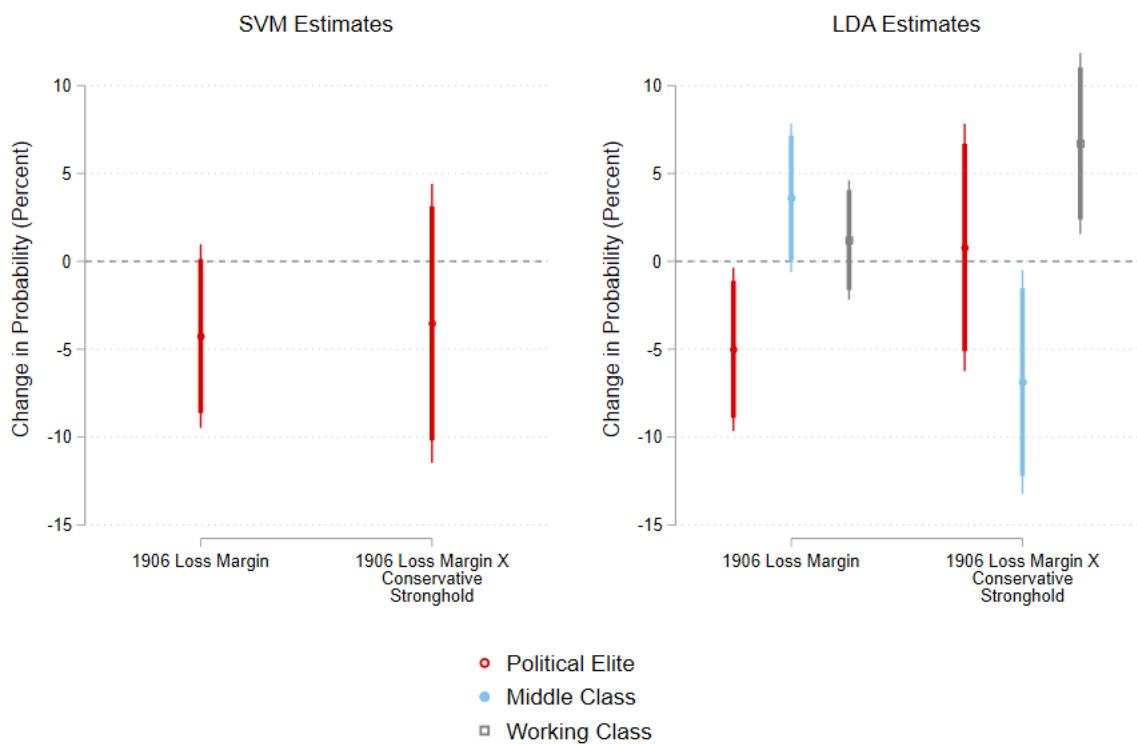
This figure illustrates the main results from Table C.3: the effect of a one standard deviation increase in the 1906 loss margin on the percentage probability that the candidate is from a certain group. Both 90 and 95 percent confidence intervals are illustrated.

I interact the 1906 loss margin with a binary variable indicating whether the Conservatives had won in the constituency four or more times in the five elections before 1906[16] (defined as a Conservative stronghold) or not. This is a good measure of how comfortable the Conservatives

---

[16]Elections before 1885 were contested on different constituency boundaries.

were in that seat, and therefore how much of a shock a given 1906 result was perceived.

The results presented in Figure 6 suggests that there was more aggressive switching of candidates in Conservative strongholds than in others. The left-hand side of Figure 6 shows the effect of the 1906 loss margin on other constituencies, and suggests switching from the political elite to the middle/working class. However, the right-hand side shows more aggressive switching: not only some evidence of switching from the political elite to working-class candidates, but also switching away from middle-class candidates. These results are consistent with a learning process where the Conservatives observe the same defeat in 1906 as a greater signal to change their candidate if they have traditionally performed well in the constituency.

In other words, the same defeat is perceived as a harsher lesson when the party felt more complacent beforehand, and as a result, the response is more radical. This result also reduces concerns about the baseline results being driven by long-term changes in voter preferences towards working-class interests - if this were the case one would not expect more aggressive switching of candidates in Conservative strongholds.

8.3 Targeting Working-Class Constituencies

In the conceptual framework in Section 2, I argue that parties tailor the portfolio of candidates as targeted instruments to win votes. If such targeting occurs, one would expect the shock of the 1906 election to have different effects on candidate choices in 1910 depending on the type of constituency they were running in. For example, if working-class constituencies have a preference for working-class candidates (whether as a commitment to a working-class platform in the future or because of some other personal affinity as explained by social identity theory), one would expect more aggressive switching to working-class candidates through a learning process.

I include an interaction between the Conservative's 1906 loss margin and a dummy variable indicating whether the constituency is described as a predominantly working-class constituency by Blewett (1972). The results are illustrated in Figure 7. While there is some evidence of switching away from political elite candidates to middle/working-class candidates in other constituencies, there is far more aggressive and significant switching in working-class constituencies. The right-hand side of Figure 7 illustrates evidence of switching to solely working-class candidates in response to a greater defeat in 1906.

26

Figure 7: Working-Class Constituencies and Candidate Switching
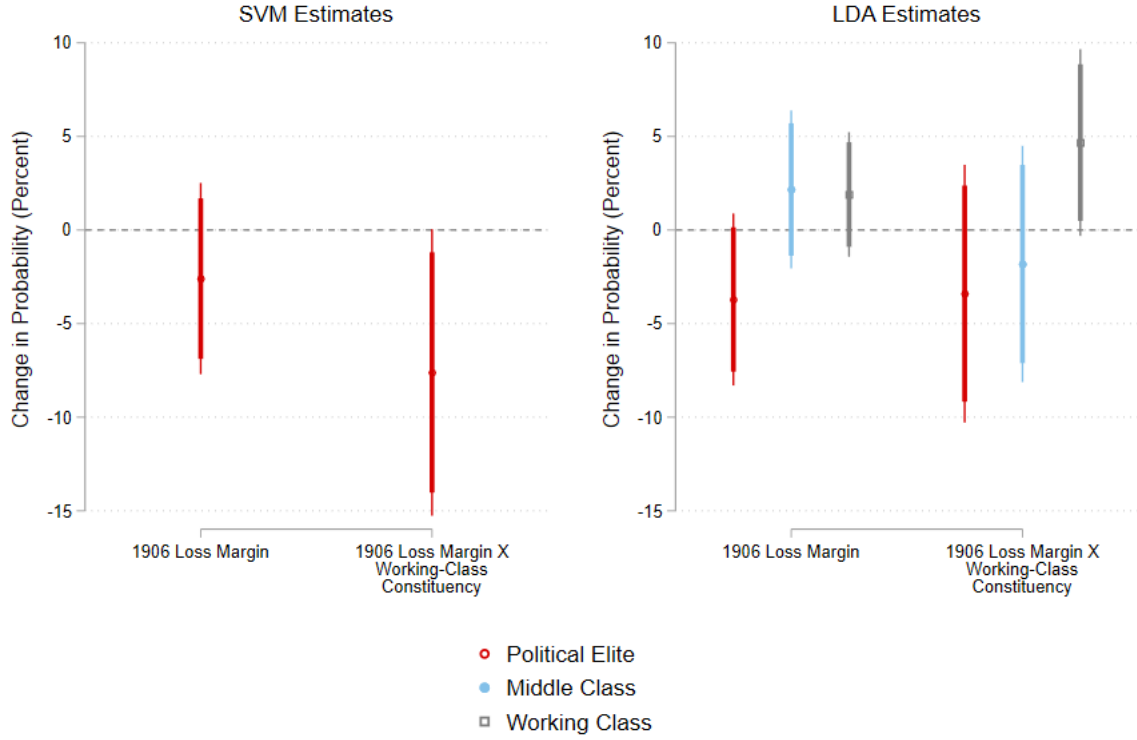
This figure illustrates the main results from Table C.4: the effect of a one standard deviation increase in the 1906 loss margin on the percentage probability that the candidate is from a certain group. Both 90 and 95 percent confidence intervals are illustrated.

9 A Winning Strategy? Implications for the 1910 Elections

The discussion of the results provide evidence of the political elite responding to the 1906 election by changing representation within the candidates in the Conservative party in 1910. One following question may be whether this strategy actually benefited them, as the previous section suggests the Conservatives may have expected by placing more working-class candidates in working-class areas. Although I do not claim a strong causal link in this analysis, the evidence is consistent with the response of targeted switching of candidates being beneficial.

I investigate whether replacing candidates aided the Conservatives in the 1910 elections. I implement a different empirical strategy which is outlined below:

$$WinMargin_{i,1910} = \alpha + \sum_{C=U,M,W} \beta^C y_{i,1910} \times ConstituencyTypeC_i \qquad (2)$$
$$+ \gamma_1 y_{i,1906} + \gamma_2 WinMargin_{i,1906} + \epsilon_{i,1910}$$

$WinMargin_{i,1910}$ is the proportion of the vote that the Conservatives won by in 1910 in con-

stituency i, while $y_{i,t}$ is the characteristic of the Conservative candidate. I differentiate the effect of changing the candidate by the type of constituency, where the type of constituency C could be upper, middle, or working class (U, M, W) as defined by Blewett (1972). As discussed in the conceptual framework, different constituencies may respond differently to different types of candidates, so I include interaction terms between the candidate characteristic and the type of constituency. I also include controls of past performance $WinMargin_{i,1906}$ and past candidate characteristics $y_{i,1906}$.

Figure 8: Selecting Working-Class Candidates and the 1910 Elections
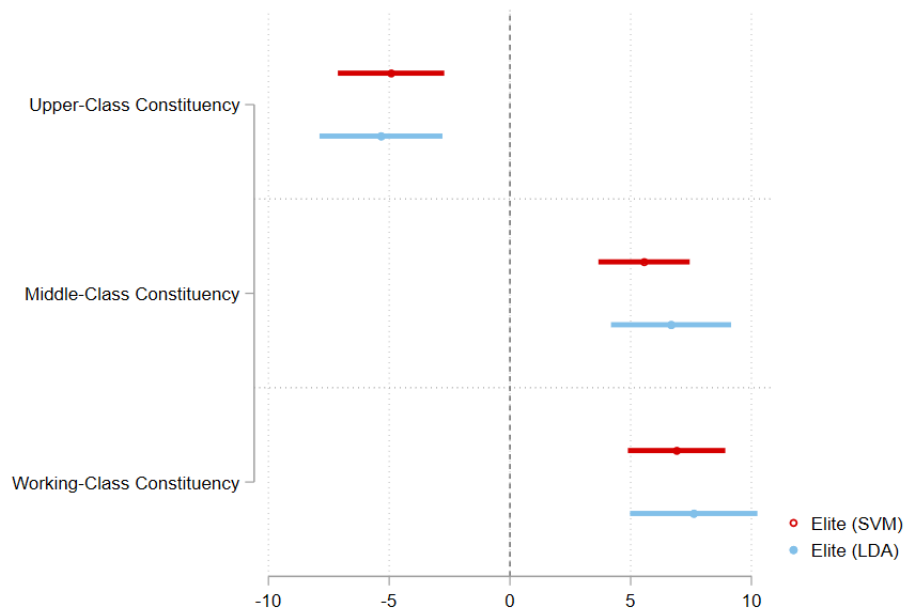
This figure illustrates the estimated impact of a one standard deviation change in the probability that the candidate is from the political elite (as estimated by the SVM/LDA) on the Conservative win margin in 1910, differentiating across constituencies.

Figure 8 illustrate the results, and the effect of the Conservatives moving from one candidate to another who is one standard deviation less likely to be a candidate from the political elite, for different types of constituencies. In upper-class constituencies, this replacement would cost the Conservatives around 5 per cent of the vote. The contrast to middle-class and working-class constituencies is significant: the Conservatives vote share increase by around 6 and 7 percentage points by this replacement, respectively.

As suggested in the previous section, the Conservatives appear to learn and understand this pattern. There are some limitations to this analysis. For example, the Liberal and Labour election pact of 1903 had somewhat broken by 1910, which because of the plurality voting system may have aided the Conservative recovery, especially in working-class constituencies. However, I argue that this is not a grave concern for two reasons. Firstly, the above analysis is

concerned with vote shares and it seems unlikely that the Conservatives would have gained votes from Liberal voters switching to Labour candidates. Secondly, Labour and the Liberals only contested 4.5 percent of constituencies across the 1910 elections, so is a relatively small problem. A greater problem is whether an increase in Conservative vote share in 1910 is likely to arise from factors that also influence the switching of candidates. While the Conservatives kept their manifesto the same, they may have made other efforts (for example in local politics) to ingratiate themselves with a broader set of voters. As a result, the coefficient estimate of the effect of candidates on vote share may overestimate the impact of the candidates themselves. However, the results reported are consistent with the targeted switching of candidates being a successful strategy.

10   Conclusion

There are many reasons why some groups are particularly under-represented in politics, and persistently so. I investigate the role of parties in this under-representation, and find evidence that their electoral considerations impact political representation. I find that losing an election can teach parties to cast the net wider for candidates in order to reach a wider range of voters.

These results have implications for the under-representation of other groups in politics, and in other occupations. While there are still persistent inequalities in political representation, these results suggests that parties can be motivated to widen political representation through losing elections. Further work would ask how persistent these effects are, and how the increased inclusivity changes policy in the short and long term. Unfortunately the much-changed political landscape of British politics at the next election in 1918 make it difficult to make undertake this analysis[17]. In addition, further work would investigate under what circumstances do political parties learn to be more inclusive. Although this paper provides an example of a party experimenting with more inclusive representation, it is important to understand under what conditions will the experiment chosen be more inclusive representation among candidates.

These results have wider implications beyond politics. They suggest that organisations may be inclusive in their membership in response to failure, whether it be in their hiring or selection of key decision makers. While organisations may be slow to learn, this paper provides some evidence that organisations may eventually learn to be more inclusive, even if only through a stop-start process of learning by losing.

---

[17]For example, some women were given the vote in 1918 for the first time as well as all men irrespective of wealth, as well as a raft of other social and political changes throughout the First World War.

References

Acemoglu, D. & Robinson, J. A. (2000), 'Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective', *The Quarterly Journal of Economics* **115**(4), 1167–1199.

Aidt, T. S., Dutta, J. & Loukoianova, E. (2006), 'Democracy comes to Europe: Franchise extension and fiscal outcomes 1830-1938', *European Economic Review* **50**(2), 249–283.

Aidt, T. S. & Jensen, P. S. (2013), 'Democratization and the size of government: Evidence from the long 19th century', *Public Choice* **157**(3-4), 511–542.

Aidt, T., Winer, S. & Zhang, P. (2020), Franchise Extension and Fiscal Structure in the United Kingdom 1820-1913: A New Test of the Redistribution Hypothesis.

Akerlof, G. A. & Kranton, R. E. (2000), 'Economics and Identity', *Quarterly Journal of Economics* **115**(3), 715–753.

Allen, R. (1994), Real Incomes in the English-Speaking World, 1879-1913, *in* G. Grantham & M. MacKinnon, eds, 'Labour Market Evolution: The Economic History of Market Integration, Wage Flexibility and Employment'.

Anastasopoulos, L. (2016), 'Estimating the gender penalty in House of Representative elections using a regression discontinuity design', *Electoral Studies* **43**, 150–157.

Ansolabehere, S. & Puy, M. S. (2016), 'Identity voting', *Public Choice* **169**(1-2), 77–95.

Ashforth, B. E. & Mael, F. (1989), 'Social Identity Theory and the Organization', *Academy of Management Review* **14**(1), 20–39.

Bandiera, O., Prat, A., Hansen, S. & Sadun, R. (2020), 'CEO Behavior and Firm Performance', *Journal of Political Economy* **128**(4).

Barnes, T. D. & Saxton, G. W. (2019), 'Working-Class Legislators and Perceptions of Representation in Latin America', *Political Research Quarterly* **72**(4), 910–928.

Beach, B., Jones, D. B., Twinam, T. & Walsh, R. (2019), Minority Representation in Local Government.

Bertocchi, G., Dimico, A., Lancia, F. & Russo, A. (2020), 'Youth Enfranchisement, Political Responsiveness, and Education Expenditure: Evidence from the US', *American Economic Journal: Economic Policy* **12**(3), 76–106.

Besley, T. & Persson, T. (2019), The Rise of Identity Politics.

Bhalotra, S. & Clots-Figueras, I. (2014), 'Health and the political agency of women', *American Economic Journal: Economic Policy* **6**(2), 164–197.

Bhalotra, S., Clots-Figueras, I. & Iyer, L. (2018), 'Pathbreakers? Women's Electoral Success and Future Political Participation', *The Economic Journal* **128**(613), 1844–1878.

Blei, D. M., Ng, A. Y. & Edu, J. B. (2003), 'Latent Dirichlet Allocation', *Journal of Machine Learning Research* **3**, 993–1022.

Blewett, N. (1972), *The Peers, the Parties, and the People: The General Election of 1910*, Palgrave Macmillan.

Boix, C. (2003), *Democracy and Redistribution*, Cambridge University Press, Cambridge.

Burden, B. (2007), *Personal Roots of Representation*, Princeton University Press, Princeton.

Carnes, N. (2012), 'Does the Numerical Underrepresentation of the Working Class in Congress Matter?', *Legislative Studies Quarterly* **37**(1), 5–34.

Carnes, N. (2016), 'Why are there so few working-class people in political office? Evidence from state legislatures', *Politics, Groups, and Identities* **4**(1), 84–109.

Carnes, N. & Hansen, E. R. (2016), 'Does paying politicians more promote economic diversity in legislatures?', *American Political Science Review* **110**(4), 699–716.

Catalano Weeks, A. (2019), 'Quotas and Party Priorities: Direct and Indirect Effects of Quota Laws', *Political Research Quarterly* **72**(4), 849–862.

Clots-Figueras, I. (2011), 'Women in politics. Evidence from the Indian States', *Journal of Public Economics* **95**(7-8), 664–690.

Clots-Figueras, I. (2012), 'Are Female Leaders Good for Education? Evidence from India', *American Economic Journal: Applied Economics* **4**(1), 212–244.

Conover, P. J. (1984), 'The Influence of Group Identifications on Political Perception and Evaluation', *The Journal of Politics* **46**(3), 760–785.

Cortes, C. & Vapnik, V. (1995), 'Support-vector networks', *Machine Learning* **20**(3), 273–297.

Craig, F. (1974), *British Parliamentary Election Results*, Macmillan Press, London.

Crone, T. (2005), 'An Alternative Definition of Economic Regions in the United States Based on Similarities in State Business Cycles', *Review of Economics and Statistics* **87**(4), 617–626.

Downs, A. (1957), 'An Economic Theory of Political Action in a Democracy', *Journal of Political Economy* **65**(2), 135–150.

Draca, M. & Schwarz, C. (2019), How Polarized are Citizens? Measuring Ideology from the Ground-Up.

Folke, O. & Rickne, J. (2016), 'The Glass Ceiling in Politics: Formalization and Empirical Tests', *Comparative Political Studies* **49**(5), 567–599.

Fulton, S. A. (2012), 'Running Backwards and in High Heels', *Political Research Quarterly* **65**(2), 303–314.

Gelpi, C. & Feaver, P. D. (2002), 'Speak softly and carry a big stick? Veterans in the political elite and the American use of force'.

Ghoddusi, H., Creamer, G. G. & Rafizadeh, N. (2019), 'Machine learning in energy economics and finance: A review', *Energy Economics* **81**, 709–727.

Gieczewski, G. (2020), Policy Persistence and Drift in Organizations.

Gieczewski, G. & Kosterina, S. (2019), Endogenous Experimentation in Organizations.

Gründler, K. & Krieger, T. (2016), 'Democracy and growth: Evidence from a machine learning indicator', *European Journal of Political Economy* **45**, 85–107.

Hainmueller, J., Hall, A. B. & Snyder Jr., J. M. (2015), 'Assessing the External Validity of Election RD Estimates: An Investigation of the Incumbency', *Source: The Journal of Politics* **77**(3), 707–720.

Hansen, S., McMahon, M. & Prat, A. (2018), 'Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach', *The Quarterly Journal of Economics* **133**(2), 801–870.

Heath, O. (2018), 'Policy Alienation, Social Alienation and Working-Class Abstention in Britain, 1964-2010', *British Journal of Political Science* **48**(4), 1053–1073.

Hopkins, D. J. & McCabe, K. T. (2012), 'After It's Too Late: Estimating the Policy Impacts of Black Mayoralties in U.S. Cities', *American Politics Research* **40**(4), 665–700.

Jensen, J. L. & Yntiso, S. (2019), 'Democratic reversals and the size of government', *European Journal of Political Economy* **58**, 21–49.

Kahn, K. F. (1996), *The political consequences of being a woman : how stereotypes influence the conduct and consequences of political campaigns*, Columbia University Press.

Keller, G., Rady, S. & Cripps, M. (2005), 'Strategic experimentation with exponential bandits'.

Kotera, G. & Okada, K. (2017), 'How does democratization affect the composition of government expenditure?', *Journal of Economic Behavior and Organization* **137**, 145–159.

Lee, D. S. (2008), 'Randomized experiments from non-random selection in U.S. House elections', *Journal of Econometrics* **142**(2), 675–697.

Lindbeck, A. & Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public Choice* **52**(3), 273–297.

Manavalan, B. & Lee, J. (2017), 'SVMQA: support-vector-machine-based protein single-model quality assessment', *Bioinformatics* **33**(16), 2496–2503.

Meltzer, A. H. & Richard, S. F. (1981), 'A Rational Theory of the Size of Government', *Journal of Political Economy* **89**(5), 914–927.

Sanbonmatsu, K. & Dolan, K. (2009), 'Do Gender Stereotypes Transcend Party?', *Political Research Quarterly* **62**(3), 485–494.

Shayo, M. (2009), 'A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution', *American Political Science Review* **103**(2), 147–174.

Stenton, M. & Lees, S. (1979), *Who's Who of British Members of Parliament: 1886-1918*, Harvester Publisher, London.

Strulovici, B. (2010), 'Learning While Voting: Determinants of Collective Experimentation', *Econometrica* **78**(3), 933–971.

Sykes, A. (1998), 'Radical Conservatism and the Working Classes in Edwardian England: The Case of the Workers Defence Union', *The English Historical Review* **113**(454), 1180–1181.

Tajfel, H. (1974), 'Social identity and intergroup behaviour', *Social Science Information* **13**(2), 65–93.

Washington, E. L. (2008), 'Female socialization: How daughters affect their legislator fathers' voting on women's issues', *American Economic Review* **98**(1), 311–332.

Wu, Y., Liu, M., Zheng, W. J., Zhao, Z. & Xu Hua (2012), Ranking Gene-Drug Relationships in Biomedical Literature Using Latent Dirichlet Allocation, *in* 'Pacific Symposium on Biocomputing', pp. 422–433.

Zhang, Y., Dong, Z., Liu, A., Wang, S., Ji, G., Zhang, Z. & Yang, J. (2015), 'Magnetic resonance brain image classification via stationary wavelet transform and generalized eigen-

value proximal support vector machine', *Journal of Medical Imaging and Health Informatics* **5**(7), 1395–1403.

Appendices

A   Definitions, Sources, and Descriptive Statistics

A.1   Description of Characteristics

Unless otherwise stated, the characteristics listed below are coded as binary variables: taking a value of one if the candidate has that attribute or characteristic, and a value of zero otherwise.

Titles

- *Honorary Title:* The candidate has been awarded an honorary title, such as an OBE, an MBE or a knighthood. This includes the investiture of peerages where the title has not been inherited from family members.

- *Hereditary Title:* The candidate has inherited a title from a family, most commonly their father or brother. This includes titles such as Earl, Lord, Baron, and Baronet.

- *Military Title:* The candidate has been given a significant military honour. This includes both the appointment to a high-ranking military position such as Colonel or General and awards for notable military service such as the Victoria Cross.

- *Title:* The written name of the candidate on the ballot includes a title beyond 'Mr.'. This includes titles such as Sir, Dr., Lord, Viscount, and Reverend.

- *Family Hereditary Title:* A member of the candidate's family holds a title that is inherited - this does not necessarily mean that the candidate themselves ever inherits this title (for example it may be held by their uncle who passes the title down to their son, the candidate's cousin).

Education

- *Eton:* The candidate attended Eton School.

- *Public School:* The candidate attended one of the schools included in the Public Schools Act 1868. The schools (Charterhouse, Eton, Harrow, Rugby, Shrewsbury, Westminster, Winchester) are among the most elite fee-paying schools.

- *Fee Paying School:* The candidate attended a school that required fees for attendance. This would not be the case for grammar or local state schools.

- *Oxbridge:* The candidate attended at least one of the University of Cambridge and the

University of Oxford.

- *University:* The candidate attended university, whether in the United Kingdom or abroad.

Occupations

- *Family Occupations:* These include the notable occupations (as mentioned in the biographies) of the candidate's family members. In most cases this will include the occupation(s) of their father, but in addition may also include the occupation(s) of other family members such as their grandparents or uncles/aunts.

- *Previous Occupations:* These include the candidate's previous occupations before running for office, and therefore may include more than one occupation. In some cases the timing of some employment is unclear so the author's judgement was used to determine whether such employment was before or after their candidacy.

Occupational Groups

- *Professionals:* These include the professions below, primarily consisting of white-collar workers or administrators.

  - *Academia:* One is defined as being part of academia if they hold an academic or teaching post, for example a Fellow or Lecturer at a university. In addition, this may include those with no formal post but are notable for their prolific publications, for example literary writers.

  - *Civil Service:* The civil service includes most appointed public office, whether in an administrative or managerial capacity. Recurring examples include diplomats or high-ranking non-partisan Whitehall advisers.

  - *Journalism:* This includes work as a journalist or editor for a newspaper. Notably this excludes those who are newspaper proprietors, who are classified as involved in business.

  - *Law:* Working in the legal profession whether as a barrister or a solicitor.

  - *Church:* Religious occupations, the vast majority of these cases are local vicars or priests.

- *Business and Finance:* This is the collective term for those who work in business or finance.

  - *Business:* This includes any management or ownership of a company from any sector/industry.

– *Finance:* This is defined as the work in the financial sector, such as investment banking.

- *Working Class:* This is the collective term for working-class or blue collar occupations.

    – *Engineering:* A catch-all term for more technical blue-collar work, such as foremen or technicians.

    – *Manual Labour:* For example manual work in factories.

    – *Trade Union:* Activity in supporting or running a trade union, whether locally or nationally.

- *Military:* Any military occupation at any level in the army or navy.

- *Politics:* Involvement in politics at either a local or national level.

    – *Local Politics:* Examples include being on the board of a local school or committee, or local council. In many cases not this work is not necessarily affiliated with a political party.

    – *Politics/Party Politics:* This includes work in party politics, whether in the organisation of a political party or in the aid of MPs at Westminster (such as an assistant).

Social Connections

- *Number of Clubs:* The number of social clubs the candidate is a member of as listed in their biographies. Social clubs played a significant role in not only the social lives of those in the upper/upper-middle class, but also their political lives.

- *Brooks, Carlton, National Liberal, Reform:* These constitute the social clubs who are most represented among politicians. For example, the National Liberal and Reform clubs include central members of the Liberal Party, while the Brooks and Carlton those of the Conservative Party. Membership is not automatic, so inclusion in one of these is a good marker of social and political status.

- *Inner Temple, Middle Temple, Lincoln's Inn, Gray's Inn:* Membership of the London law houses. The vast majority of lawyers in the sample are a member of one of these, with the exception of practising in Scotland.

- *School Connections:* The number of connections to other candidates via the school that they attended.

- *University Connections:* The number of connections to other candidates via the university that they attended.

- *Legal Connections:* The number of connections to other candidates via the law house they practised in.

- *JP:* Whether the candidate was appointed as a Justice of the Peace. A local legal occupation, but the appointment itself indicates local political influence.

A.2 Data Sources

Figure A.1: Sample Extract: Who's Who of British Members of Parliament: 1886-1918

CAS — LONDON

nkers;
ion of
he Uni-
en he
sity in
-Chair-
London
-Lieut.
enancy
British
innean
various
hor of
tion and
igin and
Flowers
Years of
Represen-
wers and
ngs, The
nd, The
is scien-
t. 1865.
nmittee
ist. Un-
865 and
1870 to
uccessful
niversity
rd Ave-
Andrews
3. [1899]

ornoway
's, Lon-
Market,
and St.
Esq., of
Lydia, d.

1895. Elected for the Lowestoft division of Suffolk in 1900 and sat until defeated in 1906. Contested Kennington in Jan. and Dec. 1910. Died 11 Dec. 1918.   [1905]

LUCAS, Reginald Jaffray. Queen Anne's Mansions, London. Carlton, Turf, Marlborough, and Bachelors'. S. of Sir Thomas Lucas, Bart. B. in London 1865. Unmarried. Educ. at Eton, and Trinity Coll., Cambridge. Private Secretary to the Rt. Hon. A. Akers-Douglas, Chief Unionist Whip 1886-92; to the Rt. Hon. Sir W.H. Walrond, Chief Unionist Whip 1895-1900. Wrote *Felix Dorrien* and other novels. Was Lieut. London Rifle Brigade, Aide-de-Camp to Brigadier commanding Fourth London Volunteer Brigade, later Capt. 3rd Battalion (M.) Hampshire Regiment. A Conservative, and a loyal supporter of the Church of England, strongly opposed to lawlessness and illegal innovations in church practice. Earnestly desired legislation for better housing of the working classes in London and other great towns. Elected for Portsmouth in 1900 and sat until he retired in Jan. 1906, and unsuccessfully contested Bury at the same time. Died 9 May 1914.   [1905]

LUCAS-SHADWELL, William. 1 Curzon Street, London. The Hall, Fairlight, Hastings. Carlton, New University, and Constitutional. Only s. of W.D. Lucas-Shadwell, Esq., Dept.-Lieut. and J.P. for Sussex, of The Hall, Fairlight, near Hastings, and Florentia M., only child and heiress of the Rev. H. Wynch, Rector of Pett, Sussex. B. 1850, at Fairlight.

Figure A.2: Sample Extract: Oxford Dictionary of National Biography

Burns, John Elliott 🔓

Previous version

Burns, John Elliott (1858–1943), labour leader and politician, was born in South Lambeth, London, on 20 October 1858, the sixteenth child of Alexander Burns, a Scottish fitter, and Barbara Smith. He left St Mary's national school when he was about ten and after a series of short-term jobs was apprenticed as an engineer.

Burns spent his childhood in relative poverty. Apparently deserted by his father, his mother took in washing and the family moved to a basement dwelling in Battersea. Despite his brief schooling Burns attended night classes and became an avid consumer of radical literature. A useful boxer, he was a pugnacious and combative individual, and insubordination twice resulted in the cancellation of his indentures by irate employers before he finally completed his

John Elliott Burns (1858–1943)
by John Collier, 1889
© National Portrait Gallery, London

apprenticeship at Mowlems, a major London contractor. There he came under the influence of Victor Delahaye, an exiled former communard and member of a French Marxist group, committed to the view that in the absence of effective physical force and organization, the working class could throw off capitalism only through the ballot box. In 1878 Burns had his first encounter with the law when he was arrested for holding a

A.3 Election Results 1900-1910

Figure A.3: Conservative Margin of Defeat in the 1906 UK Election
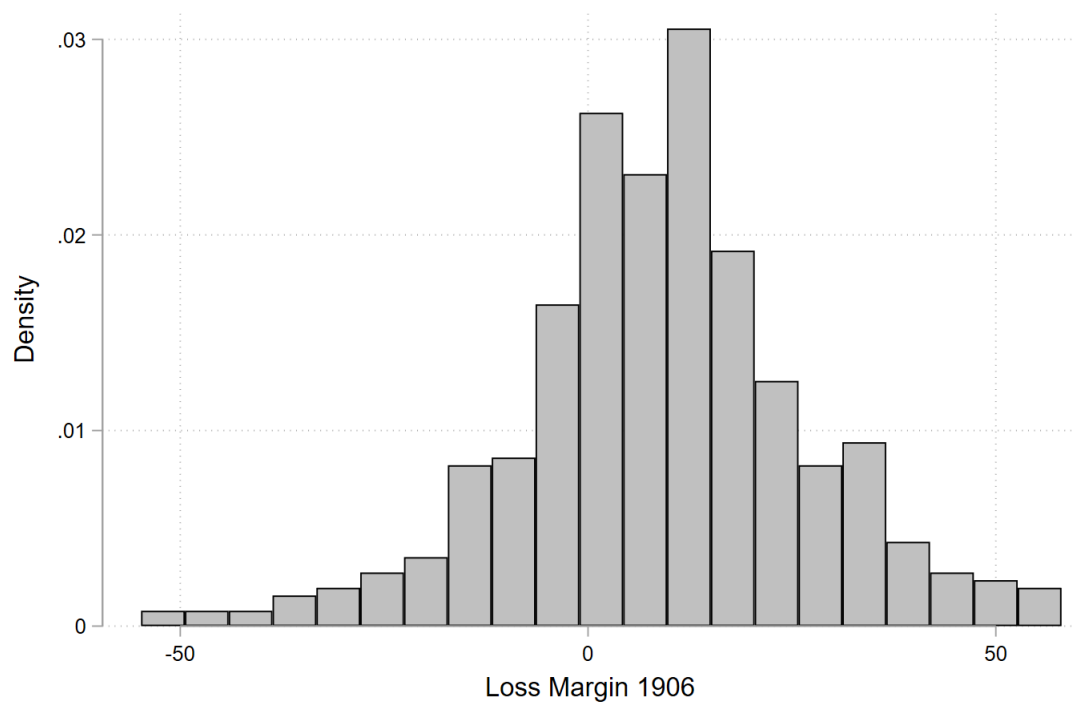
This figure illustrates the distribution of the Conservative loss margin (the proportion of the vote won by the Conservatives minus the proportion of the vote won by the most successful other party in the constituency) across constituencies in 1906.

Figure A.4: Composition of Parliament from the 1900 to 1910 (Dec.) General Elections
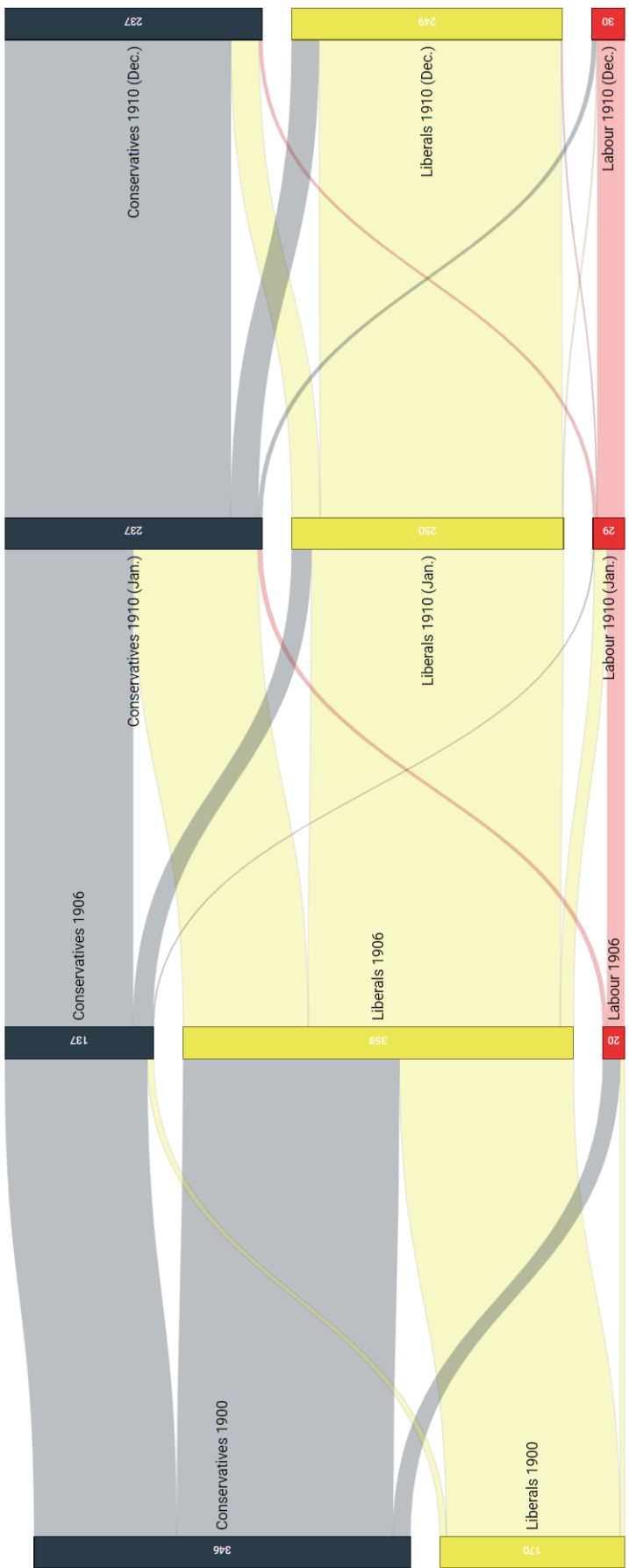
A.3.1 Education, Titles, and Family Backgrounds

Table A.1: Summary Statistics: Education, Titles and Family Backgrounds (All Candidates)

	(1) Whole Sample mean	(2) 1906 (Jan.) mean	(3) 1910 (Dec.) mean	(4) Difference mean
Honorary Title	48.9	49.4	44.9	-9.2
Hereditary Title	20.0	21.8	16.0	-26.3
Military Title	10.4	9.8	10.0	2.3
Title	22.0	23.0	20.1	-12.9
Family Hereditary Title	25.7	26.2	23.9	-8.7
Eton	15.5	16.3	14.4	-11.4
Public School	28.0	29.3	26.0	-11.2
Fee Paying School	81.3	81.2	79.9	-1.6
Oxbridge	34.7	35.5	32.1	-9.5
University	51.4	52.4	48.7	-7.1
Professional (Family)	28.8	30.4	27.6	-9.4
Academia (Family)	9.1	9.9	9.3	-6.3
Civil Service (Family)	3.8	3.9	3.9	-2.1
Journalism (Family)	0.4	0.6	0.3	-45.8
Law (Family)	6.2	6.3	6.3	0.1
Church (Family)	12.8	13.5	11.5	-14.8
Business and Finance (Family)	17.7	17.6	16.7	-5.2
Business (Family)	14.5	14.1	13.8	-1.9
Finance (Family)	3.8	4.2	3.1	-26.0
Working Class (Family)	10.1	10.4	10.6	1.4
Engineering (Family)	2.2	2.1	2.2	3.5
Labourer (Family)	8.1	8.5	8.6	1.1
Trade Union (Family)	0.0	0.0	0.0	.
Military (Family)	14.4	14.5	14.4	-0.9
Politics and Local Politics (Family)	27.8	28.6	26.4	-7.6
Local Politics (Family)	15.1	15.4	14.9	-3.1
Politics (Family)	13.9	14.5	12.3	-14.7
Observations	3938	1038	957	905

This table reports summary statistics for a range of characteristics for all candidates in the sample, which are explained in more detail in Appendix A.1. Columns 2 and 3 focus on the the 1906 and December 1910 (the last election in the sample) elections, and Column 4 reports the percentage change in the proportion from 1906 to December 1910.

Table A.2: Summary Statistics: Education, Titles and Family Backgrounds (Conservative Candidates)

	(1) Whole Sample mean	(2) 1906 (Jan.) mean	(3) 1910 (Dec.) mean	(4) Difference mean
Honorary Title	50.9	54.5	45.1	-17.2
Hereditary Title	29.3	34.3	23.0	-32.7
Military Title	14.7	14.3	14.2	-0.3
Title	25.3	32.4	18.3	-43.4
Family Hereditary Title	34.0	35.0	31.2	-11.1
Eton	24.5	26.4	23.2	-12.4
Public School	40.0	42.2	37.1	-12.2
Fee Paying School	91.6	91.1	91.3	0.2
Oxbridge	41.4	41.0	39.2	-4.5
University	55.2	54.7	53.1	-3.0
Professional (Family)	28.4	29.9	27.6	-7.8
Academia (Family)	8.7	9.4	9.5	0.5
Civil Service (Family)	3.8	3.3	3.8	15.6
Journalism (Family)	0.4	0.6	0.2	-65.8
Law (Family)	6.0	6.1	6.5	6.2
Church (Family)	12.9	13.7	11.8	-14.1
Business and Finance (Family)	14.2	14.3	12.2	-14.9
Business (Family)	11.4	11.1	10.1	-8.7
Finance (Family)	3.5	3.9	2.3	-40.5
Working Class (Family)	3.3	2.7	3.6	34.3
Engineering (Family)	1.8	1.6	1.7	2.7
Labourer (Family)	1.5	1.0	1.9	84.9
Trade Union (Family)	0.0	0.0	0.0	.
Military (Family)	19.4	18.9	20.2	7.2
Politics and Local Politics (Family)	27.5	28.3	24.6	-12.9
Local Politics (Family)	15.2	15.4	14.5	-5.5
Politics (Family)	14.0	14.5	11.4	-21.9
Observations	1973	488	475	459

This table reports summary statistics for a range of characteristics for Conservative candidates in the sample, which are explained in more detail in Appendix A.1. Columns 2 and 3 focus on the the 1906 and December 1910 (the last election in the sample) elections, and Column 4 reports the percentage change in the proportion from 1906 to December 1910.

A.3.2 Previous Occupations and Social Connections

Table A.3: Summary Statistics: Previous Occupations and Social Connections (All Candidates)

	(1) Whole Sample mean	(2) 1906 (Jan.) mean	(3) 1910 (Dec.) mean	(4) Difference mean
Professional	55.5	56.3	52.2	-7.1
Academia	15.7	17.1	14.4	-15.9
Civil Service	32.5	33.6	27.8	-17.3
Journalism	3.5	4.1	3.6	-14.2
Law	24.7	24.0	24.2	1.1
Church	1.3	1.5	1.1	-25.4
Business and Finance	29.4	29.7	27.9	-6.0
Business	26.2	26.2	24.7	-5.9
Finance	5.5	5.6	5.5	-0.9
Working Class	8.0	8.5	8.3	-2.6
Engineer	2.4	2.2	2.5	13.2
Labourer	5.1	5.8	5.0	-13.2
Trade Union	5.5	6.4	5.9	-8.0
Military	24.4	24.8	23.7	-4.2
Politics and Local Politics	36.0	36.6	34.5	-5.8
Local Politics	34.9	35.6	33.0	-7.4
Politics	1.5	1.4	2.2	51.8
Number of Clubs	1.6	1.7	1.5	-6.5
Athenaeum Club	6.4	6.8	5.4	-20.6
Brooks Club	7.6	8.3	6.6	-20.5
Carlton Club	29.1	27.5	27.9	1.6
National Liberal Club	17.0	18.8	16.1	-14.3
Reform Club	17.1	18.8	15.2	-19.3
Inner Temple	8.5	8.6	7.6	-11.0
Middle Temple	4.3	4.0	4.5	11.0
Lincoln's Inn	3.4	3.9	2.9	-24.1
Gray's Inn'	0.8	0.3	1.1	297.7
School Connections	61.3	64.0	57.2	-10.6
University Connections	135.2	138.4	126.2	-8.8
Legal Connections	20.9	21.0	19.2	-8.6
JP	36.6	37.9	31.6	-16.8
Observations	3938	1038	957	905

This table reports summary statistics for a range of characteristics for all candidates in the sample, which are explained in more detail in Appendix A.1. Columns 2 and 3 focus on the the 1906 and December 1910 (the last election in the sample) elections, and Column 4 reports the percentage change in the proportion from 1906 to December 1910.

Table A.4: Summary Statistics: Previous Occupations and Social Connections (Conservative Candidates)

	(1) Whole Sample mean	(2) 1906 (Jan.) mean	(3) 1910 (Dec.) mean	(4) Difference mean
Professional	52.3	53.5	48.0	-10.3
Academia	11.1	12.7	9.5	-25.4
Civil Service	30.9	33.2	24.6	-25.8
Journalism	1.9	2.9	1.7	-41.3
Law	24.6	22.7	25.1	10.1
Church	0.9	1.0	0.4	-58.9
Business and Finance	24.8	25.2	21.5	-14.8
Business	21.4	21.3	18.3	-14.1
Finance	5.0	4.9	4.4	-10.1
Working Class	2.6	1.4	2.9	105.5
Engineer	1.7	1.0	1.5	43.8
Labourer	0.9	0.4	1.3	208.2
Trade Union	0.7	0.4	0.8	105.5
Military	35.6	36.7	34.5	-5.9
Politics and Local Politics	35.4	36.7	31.2	-15.1
Local Politics	34.1	35.2	29.7	-15.8
Politics	1.9	2.0	2.3	13.0
Number of Clubs	1.8	1.8	1.8	-4.2
Athenaeum Club	7.3	8.4	5.5	-34.8
Brooks Club	4.7	4.9	3.8	-22.9
Carlton Club	56.6	56.8	54.3	-4.3
National Liberal Club	0.7	0.8	0.8	2.7
Reform Club	1.6	1.8	0.4	-77.2
Inner Temple	10.2	9.6	9.3	-3.8
Middle Temple	3.3	3.1	4.0	30.1
Lincoln's Inn	3.3	3.9	2.7	-29.7
Gray's Inn'	0.7	0.2	1.3	516.4
School Connections	92.8	98.6	87.6	-11.2
University Connections	158.4	157.9	150.7	-4.6
Legal Connections	22.8	22.0	21.5	-2.6
JP	37.6	39.1	31.1	-20.5
Observations	1973	488	475	459

This table reports summary statistics for a range of characteristics for Conservative candidates in the sample, which are explained in more detail in Appendix A.1. Columns 2 and 3 focus on the the 1906 and December 1910 (the last election in the sample) elections, and Column 4 reports the percentage change in the proportion from 1906 to December 1910.

B   Classification: Aggregation Methods

This section illustrates the correlation between different characteristics and the group classi-
fication that the two machine learning sorting algorithms undertake. Correlations between
characteristics and membership of different groups match prior expectations in the vast ma-
jority of cases.

B.1   Support Vector Machine (SVM)

Table B.1: Characteristics of the Elite Index (Support Vector Machine)

	Elite Index (SVM)		Elite Index (SVM)
		Politics	0.00388
Academia (Family)	0.0201		(0.0160)
	(0.0298)	Church	0.00209
Business (Family)	-0.140***		(0.00584)
	(0.0313)	Trade Union	-0.0489***
Civil Service (Family)	0.000751		(0.00874)
	(0.0190)	Oxbridge	0.692***
Engineer (Family)	-0.0300**		(0.0428)
	(0.0137)	University	0.569***
Finance (Family)	-0.0119		(0.0436)
	(0.0160)	Eton	0.540***
Journalism (Family)	0.00215		(0.0396)
	(0.00477)	Public School	0.844***
Law (Family)	-0.0379		(0.0393)
	(0.0250)	Published Author	0.000672*
Local Politics (Family)	0.0307		(0.000394)
	(0.0370)	Athenaeum Club	0.0765***
Military (Family)	0.0676		(0.0249)
	(0.0413)	Brooks Club	0.0909***
Church (Family)	0.0370		(0.0206)
	(0.0339)	Carlton Club	0.284***
Academia	0.000468		(0.0466)
	(0.0298)	National Liberal Club	-0.0216**
Business	-0.386***		(0.00887)
	(0.0382)	Reform Club	-0.00628
Civil Service	0.0588		(0.00824)
	(0.0444)	Number of Clubs	0.0120***
Engineer	-0.0525***		(0.00129)
	(0.0133)	School Connections	2.014***
Finance	-0.0135		(0.113)
	(0.0225)	University Connections	2.645***
Journalism	0.0110		(0.157)
	(0.0142)	Legal Connections	0.308***
Labourer	-0.0690***		(0.0537)
	(0.0103)	Honorary Title	0.00891
Law	0.0921**		(0.0490)
	(0.0448)	Hereditary Title	0.349***
Local Politics	-0.130***		(0.0407)
	(0.0466)	Military Title	0.168***
Military	0.438***		(0.0348)
	(0.0459)	JP	0.0893**
			(0.0451)

This table illustrates the characteristics of the elite index, reporting the coefficient results of simply regressing the
characteristic variable on the elite index created by the Support Vector Machine. Standard errors are reported in
parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10
percent level, respectively.

B.2 Latent Dirichlet Allocation (LDA)

Table B.2: Characteristics of the Elite Index (Latent Dirichlet Allocation)

	Elite Index (LDA)		Elite Index (LDA)
Academia (Family)	-0.169***	Politics	0.0286*
	(0.0302)		(0.0164)
Business (Family)	-0.231***	Church	-0.00438
	(0.0317)		(0.00601)
Civil Service (Family)	-0.0105	Trade Union	-0.0330***
	(0.0196)		(0.00910)
Engineer (Family)	-0.0583***	Oxbridge	0.297***
	(0.0141)		(0.0494)
Finance (Family)	-0.0262	University	0.0654
	(0.0164)		(0.0492)
Journalism (Family)	-0.00917*	Eton	0.727***
	(0.00490)		(0.0375)
Law (Family)	-0.160***	Public School	1.032***
	(0.0252)		(0.0357)
Local Politics (Family)	-0.0354	Published Author	-0.00124***
	(0.0381)		(0.000404)
Military (Family)	0.222***	Athenaeum Club	-0.0910***
	(0.0420)		(0.0256)
Church (Family)	-0.101***	Brooks Club	0.110***
	(0.0348)		(0.0211)
Academia	-0.220***	Carlton Club	0.260***
	(0.0298)		(0.0483)
Business	-0.364***	National Liberal Club	-0.0263***
	(0.0398)		(0.00912)
Civil Service	0.0256	Reform Club	-0.0256***
	(0.0458)		(0.00845)
Engineer	-0.0594***	Number of Clubs	0.00841***
	(0.0136)		(0.00136)
Finance	0.00702	School Connections	2.589***
	(0.0232)		(0.104)
Journalism	-0.0333**	University Connections	1.020***
	(0.0146)		(0.184)
Labourer	-0.0451***	Legal Connections	-0.354***
	(0.0108)		(0.0550)
Law	-0.561***	Honorary Title	-0.178***
	(0.0421)		(0.0501)
Local Politics	-0.124***	Hereditary Title	0.486***
	(0.0480)		(0.0403)
Military	0.817***	Military Title	0.335***
	(0.0411)		(0.0345)
		JP	0.0750
			(0.0466)

This table illustrates the characteristics of the political elite group as defined by Latent Dirichlet Allocation, reporting the coefficient results of simply regressing the characteristic variable on the probability that the Latent Dirichlet Allocation defines the candidate as being part of the political elite. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

Table B.3: Characteristics of the Working-Class/Political Outsider Group (Latent Dirichlet Allocation)

	Working (LDA)		Working (LDA)
		Politics	-0.0232
Academia (Family)	-0.00504		(0.0227)
	(0.0425)	Church	-0.00876
Business (Family)	0.494***		(0.00831)
	(0.0419)	Trade Union	0.0973***
Civil Service (Family)	-0.0457*		(0.0122)
	(0.0270)	Oxbridge	-0.986***
Engineer (Family)	0.101***		(0.0610)
	(0.0193)	University	-0.826***
Finance (Family)	0.0801***		(0.0619)
	(0.0226)	Eton	-0.634***
Journalism (Family)	0.0119*		(0.0583)
	(0.00678)	Public School	-1.004***
Law (Family)	-0.108***		(0.0604)
	(0.0354)	Published Author	-0.00142**
Local Politics (Family)	0.00505		(0.000560)
	(0.0527)	Athenaeum Club	-0.161***
Military (Family)	-0.170***		(0.0352)
	(0.0587)	Brooks Club	-0.105***
Church (Family)	-0.0408		(0.0295)
	(0.0483)	Carlton Club	-0.229***
Academia	-0.0445		(0.0674)
	(0.0425)	National Liberal Club	0.0531***
Business	0.929***		(0.0125)
	(0.0480)	Reform Club	0.0122
Civil Service	0.0258		(0.0117)
	(0.0633)	Number of Clubs	-0.0104***
Engineer	0.149***		(0.00189)
	(0.0184)	School Connections	-2.363***
Finance	0.110***		(0.171)
	(0.0318)	University Connections	-3.697***
Journalism	0.0514**		(0.225)
	(0.0201)	Legal Connections	-0.579***
Labourer	0.118***		(0.0753)
	(0.0146)	Honorary Title	-0.00363
Law	-0.489***		(0.0698)
	(0.0617)	Hereditary Title	-0.325***
Local Politics	0.434***		(0.0593)
	(0.0649)	Military Title	-0.218***
Military	-0.517***		(0.0497)
	(0.0664)	JP	-0.0497
			(0.0644)

This table illustrates the characteristics of the working-class/political outsiders group as defined by Latent Dirichlet Allocation, reporting the coefficient results of simply regressing the characteristic variable on the probability that the Latent Dirichlet Allocation defines the candidate as being part of the working class/political outsiders. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

C SVM/LDA Variables: Regression Tables and Figures

C.1 Falsification Test

Table C.1: Comparing Political Parties

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
Other Parties × 1906 Loss Margin	0.035	0.005	-0.015	0.016
	(0.081)	(0.063)	(0.063)	(0.059)
Conservatives × 1906 Loss Margin	-0.271**	-0.216**	0.057	0.155*
	(0.121)	(0.094)	(0.094)	(0.087)
Adjusted R^2	0.303	0.376	0.237	0.413
N	1038	1040	1040	1040

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate select in 1910, across different political parties (not just the Conservatives). The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the candidate is from Other Parties (besides the Conservatives) and another interaction with whether candidate is from the Conservatives. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

C.2 Don't Lose, Don't Learn?

Table C.2: Incumbents and Losing Candidates

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	0.188	0.140	-0.071	-0.087
	(0.267)	(0.239)	(0.220)	(0.174)
1906 Conservative Defeat $\times$ 1906 Loss Margin	-0.649**	-0.577**	0.213	0.400**
	(0.298)	(0.268)	(0.247)	(0.194)
1906 Conservative Defeat	-0.639	1.911	-1.282	-0.886
	(5.636)	(5.031)	(4.640)	(3.667)
Adjusted R^2	0.096	0.212	0.106	0.168
N	468	468	468	468

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different outcomes in 1906. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin and the 1906 loss margin interacted with whether the Conservatives lost in 1906. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). The lagged dependent variables are the values of the dependent variables in 1906. The 1900 loss margin and lagged dependent variable are also included, as well as a dummy variable that takes value of one if the Conservatives lost in that constituency in 1906. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

C.3 Conservative Strongholds

Table C.3: Conservative Strongholds

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.209	-0.247**	0.178*	0.060
	(0.131)	(0.117)	(0.106)	(0.085)
Conservative Stronghold × 1906 Loss Margin	-0.174	0.039	-0.338**	0.331**
	(0.199)	(0.177)	(0.160)	(0.129)
Adjusted R^2	0.087	0.205	0.123	0.171
N	468	468	468	468

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different types of constituencies. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the constituency is a Conservative stronghold and the 1906 loss margin interacted with whether the constituency is one of the other constituencies. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). The lagged dependent variables are the values of the dependent variables in 1906. The 1900 loss margin and lagged dependent variable are also included. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

C.4 Working-Class Constituencies

Table C.4: Working Class Constituencies

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.074	-0.162	0.094	0.082
	(0.123)	(0.111)	(0.102)	(0.080)
Working Class $\times$ 1906 Loss Margin	-0.478***	-0.205	-0.056	0.245**
	(0.165)	(0.149)	(0.137)	(0.108)
Adjusted R^2	0.103	0.206	0.106	0.169
N	468	468	468	468

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different types of constituencies. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the constituency is a defined as a working-class constituency and the 1906 loss margin interacted with whether the constituency is one of the other constituencies. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). The lagged dependent variables are the values of the dependent variables in 1906. The 1900 loss margin and lagged dependent variable are also included. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

C.5   A Winning Strategy? Implications for the 1910 Elections

Table C.5: Baseline Results

	(1) Win Margin 1910	(2) Win Margin 1910
Elite (SVM) × Upper-Class Constituency	4.923***	
	(1.120)	
Elite (SVM) × Middle-Class Constituency	-5.554***	
	(0.961)	
Elite (SVM) × Working-Class Constituency	-6.902***	
	(1.025)	
Elite (LDA) × Upper-Class Constituency		5.340***
		(1.294)
Elite (LDA) × Middle-Class Constituency		-6.672***
		(1.264)
Elite (LDA) × Working-Class Constituency		-7.611***
		(1.341)
Adjusted R^2	0.391	0.377
N	348	348

This table reports the effect of increasing the probability that a candidate is a member of the political elite by one standard deviation on the Conservative win margin, as outlined in Section 9. The signs are different to those of Figure 8, as Figure 8 illustrates the effect of decreasing the probability that a candidate is a member of the political elite, rather than increasing.

D All Characteristics: Regression Tables and Figures

D.1 Baseline Results

Figure D.1: Baseline Results
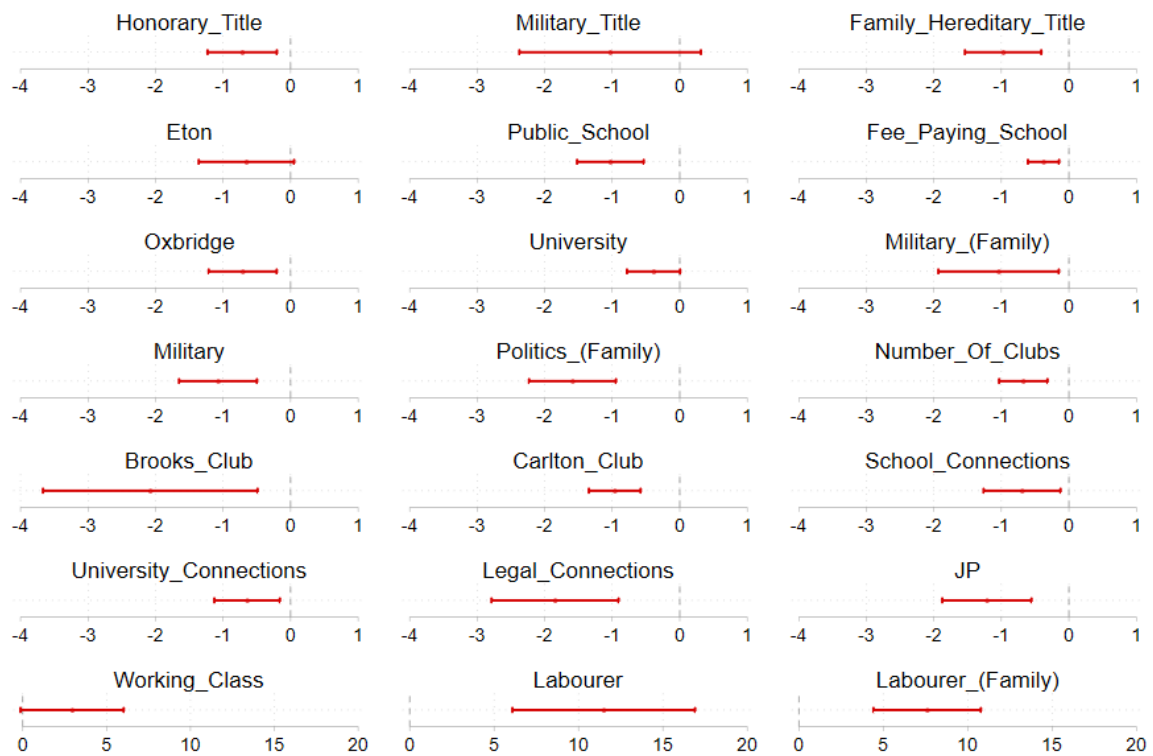
This figure illustrates the main results from Table D.1, the coefficient (and 95 per cent confidence intervals) of the 1906 loss margin on whether the Conservative candidate in 1910 has a given characteristic.

Table D.1: Baseline Results

	Loss 1906	Loss 1900	Lag. D.V.
Honorary Title	-0.716***	-0.269	0.344***
	(0.261)	(0.253)	(0.0429)
Military Title	-1.031	1.023	0.423***
	(0.684)	(0.668)	(0.0465)
Family Hereditary Title	-0.974***	-0.371	0.293***
	(0.288)	(0.312)	(0.0357)
Eton	-0.654*	-0.203	0.336***
	(0.360)	(0.391)	(0.0363)
Public School	-1.027***	-0.348	0.316***
	(0.251)	(0.276)	(0.0357)
Fee Paying School	-0.376***	0.132	0.214***
	(0.117)	(0.103)	(0.0519)
Oxbridge	-0.709***	-0.284	0.256***
	(0.256)	(0.283)	(0.0379)
University	-0.388*	-0.480**	0.259***
	(0.201)	(0.220)	(0.0384)
Military (Family)	-1.040**	0.528	0.297***
	(0.454)	(0.491)	(0.0388)
Military	-1.075***	0.178	0.259***
	(0.294)	(0.316)	(0.0376)
Politics (Family)	-1.588***	0.140	0.260***
	(0.328)	(0.354)	(0.0362)
Number Of Clubs	-0.675***	-0.425**	0.243***
	(0.183)	(0.198)	(0.0406)
Brooks Club	-2.078**	-1.028	0.285***
	(0.809)	(0.871)	(0.0309)
Carlton Club	-0.962***	-0.309	0.192***
	(0.195)	(0.216)	(0.0391)
School Connections	-0.693**	-0.158	0.365***
	(0.290)	(0.317)	(0.0362)
University Connections	-0.644***	-0.301	0.283***
	(0.247)	(0.272)	(0.0385)
Legal Connections	-1.846***	0.410	0.339***
	(0.479)	(0.517)	(0.0370)
JP	-1.214***	0.416	0.232***
	(0.336)	(0.328)	(0.0424)
Working Class	2.941*	-0.273	0.0662
	(1.567)	(1.695)	(0.0568)
Labourer (Family)	7.607***	-2.458	-0.0154
	(1.619)	(1.749)	(0.0420)
Labourer	11.50***	-3.927	0.238***
	(2.756)	(2.982)	(0.0544)

D.2 Falsification Test

Table D.2: Baseline Results

	Other × Margin 1906	Cons. × Margin 1906	Loss Margin 1900	Lagged D.V.
Honorary Title	-0.200	-0.487	-0.123	0.491***
	(0.177)	(0.304)	(0.149)	(0.0270)
Military Title	-0.0633	0.0723	-0.392	0.451***
	(0.566)	(0.973)	(0.477)	(0.0278)
Family Hereditary Title	-0.924***	0.131	-0.778***	0.500***
	(0.297)	(0.465)	(0.241)	(0.0235)
Eton	-1.038**	0.693	-0.794*	0.319***
	(0.508)	(0.796)	(0.412)	(0.0248)
Public School	-0.139	-0.990**	-0.320	0.461***
	(0.309)	(0.485)	(0.250)	(0.0253)
Fee Paying School	0.0798	-0.358**	0.00932	0.729***
	(0.0905)	(0.169)	(0.0757)	(0.0250)
Oxbridge	-0.0368	-0.562	-0.180	0.464***
	(0.229)	(0.360)	(0.186)	(0.0241)
University	-0.260	-0.121	-0.222	0.419***
	(0.167)	(0.262)	(0.135)	(0.0258)
Military (Family)	0.463	-0.866	-0.0207	0.461***
	(0.421)	(0.661)	(0.341)	(0.0233)
Military	-0.476	-0.0552	-0.201	0.516***
	(0.330)	(0.519)	(0.266)	(0.0234)
Politics (Family)	-0.183	-1.124***	-0.191	0.417***
	(0.265)	(0.415)	(0.214)	(0.0246)
Number Of Clubs	-0.255*	-0.251	-0.255**	0.546***
	(0.151)	(0.237)	(0.122)	(0.0248)
Brooks Club	-1.148**	-0.627	-1.475***	0.372***
	(0.581)	(0.914)	(0.471)	(0.0220)
Carlton Club	1.444	-3.824**	1.171	0.461***
	(1.231)	(1.924)	(0.992)	(0.0503)
School Connections	-0.539	-0.0297	-0.639**	0.343***
	(0.364)	(0.572)	(0.295)	(0.0247)
University Connections	-0.0811	-0.471	-0.202	0.445***
	(0.214)	(0.337)	(0.174)	(0.0247)
Legal Connections	-0.0213	-1.266**	0.280	0.443***
	(0.397)	(0.625)	(0.322)	(0.0245)
JP	-0.0931	-0.842**	0.255	0.452***
	(0.233)	(0.401)	(0.196)	(0.0264)
Working Class	-0.761	4.071**	-0.543	0.268***
	(1.033)	(1.621)	(0.837)	(0.0337)
Labourer (Family)	-1.826*	7.393***	-0.162	0.0900***
	(1.065)	(1.673)	(0.863)	(0.0299)
Labourer	-2.525	12.59***	-1.137	0.290***
	(1.739)	(2.731)	(1.410)	(0.0348)

D.3 Don't Lose, Don't Learn?

Table D.3: Baseline Results

	Loss Margin 1906	Defeat	Defeat × Margin	Loss Margin 1900	Lagged D.V.
Honorary Title	-0.0595	-18.18	-0.613	-0.181	0.339***
	(0.708)	(14.99)	(0.791)	(0.262)	(0.0431)
Military Title	1.227	35.45	-3.695*	0.943	0.422***
	(1.860)	(39.14)	(2.074)	(0.691)	(0.0464)
Family Hereditary Title	0.236	-11.45	-1.517*	-0.257	0.289***
	(0.793)	(19.10)	(0.891)	(0.326)	(0.0357)
Eton	1.911*	-13.24	-3.372***	0.00540	0.333***
	(0.982)	(23.69)	(1.103)	(0.407)	(0.0361)
Public School	-0.187	6.805	-1.263	-0.322	0.310***
	(0.691)	(16.68)	(0.777)	(0.288)	(0.0358)
Fee Paying School	-0.271	8.977	-0.317	0.101	0.208***
	(0.273)	(5.993)	(0.318)	(0.108)	(0.0523)
Oxbridge	-0.189	-6.431	-0.630	-0.231	0.255***
	(0.707)	(17.05)	(0.794)	(0.296)	(0.0380)
University	-0.0587	-10.68	-0.305	-0.423*	0.260***
	(0.552)	(13.34)	(0.621)	(0.231)	(0.0385)
Military (Family)	-1.022	11.68	-0.192	0.492	0.296***
	(1.252)	(30.22)	(1.407)	(0.515)	(0.0389)
Military	-0.481	49.16**	-1.540*	0.0642	0.241***
	(0.797)	(19.24)	(0.900)	(0.328)	(0.0376)
Politics (Family)	0.116	-21.40	-2.062**	0.318	0.257***
	(0.898)	(21.71)	(1.010)	(0.370)	(0.0362)
Number Of Clubs	-0.752	-1.157	0.124	-0.426**	0.244***
	(0.501)	(12.08)	(0.564)	(0.207)	(0.0408)
Brooks Club	-1.913	-100.0*	1.173	-0.702	0.279***
	(2.223)	(53.89)	(2.502)	(0.912)	(0.0311)
Carlton Club	-0.680	8.286	-0.513	-0.328	0.182***
	(0.537)	(13.03)	(0.606)	(0.227)	(0.0398)
School Connections	1.238	-3.807	-2.627***	-0.0245	0.360***
	(0.793)	(19.12)	(0.891)	(0.329)	(0.0361)
University Connections	-0.180	-5.764	-0.562	-0.253	0.282***
	(0.681)	(16.45)	(0.765)	(0.285)	(0.0387)
Legal Connections	-2.350*	-13.11	0.886	0.418	0.341***
	(1.317)	(31.77)	(1.481)	(0.542)	(0.0371)
JP	-1.119	0.326	-0.137	0.417	0.232***
	(0.929)	(19.77)	(1.037)	(0.340)	(0.0428)
Working Class	-7.829*	50.01	14.23***	-1.153	0.0698
	(4.287)	(103.5)	(4.817)	(1.765)	(0.0565)
Labourer (Family)	-0.847	-60.74	12.58**	-2.832	-0.0130
	(4.431)	(107.0)	(4.977)	(1.823)	(0.0418)
Labourer	-1.365	-74.13	18.89**	-4.561	0.240***
	(7.556)	(182.3)	(8.488)	(3.112)	(0.0542)

D.4 Conservative Strongholds

Table D.4: Baseline Results

	Other × Loss Margin 1906	Stronghold × Margin	Loss Margin 1900	Lagged D.V.
Honorary Title	-0.355	-0.601	-0.372	0.344***
	(0.368)	(0.431)	(0.263)	(0.0428)
Military Title	-1.905**	1.456	1.278*	0.420***
	(0.965)	(1.136)	(0.696)	(0.0466)
Family Hereditary Title	-0.672*	-0.746	-0.498	0.295***
	(0.344)	(0.466)	(0.321)	(0.0356)
Eton	-0.837*	0.454	-0.124	0.337***
	(0.429)	(0.583)	(0.404)	(0.0363)
Public School	-0.834***	-0.478	-0.430	0.316***
	(0.300)	(0.408)	(0.285)	(0.0357)
Fee Paying School	-0.162	-0.357*	0.0781	0.223***
	(0.163)	(0.190)	(0.107)	(0.0518)
Oxbridge	-0.480	-0.567	-0.378	0.258***
	(0.306)	(0.416)	(0.291)	(0.0379)
University	-0.105	-0.699**	-0.598***	0.261***
	(0.239)	(0.324)	(0.226)	(0.0383)
Military (Family)	-1.782***	1.837**	0.843*	0.298***
	(0.541)	(0.734)	(0.505)	(0.0386)
Military	-1.243***	0.416	0.249	0.259***
	(0.351)	(0.474)	(0.326)	(0.0376)
Politics (Family)	-1.577***	-0.0259	0.135	0.260***
	(0.391)	(0.531)	(0.366)	(0.0362)
Number Of Clubs	-0.501**	-0.427	-0.497**	0.245***
	(0.218)	(0.295)	(0.204)	(0.0406)
Brooks Club	-1.364	-1.771	-1.331	0.284***
	(0.966)	(1.308)	(0.899)	(0.0309)
Carlton Club	-1.017***	0.133	-0.288	0.190***
	(0.234)	(0.318)	(0.222)	(0.0393)
School Connections	-0.804**	0.275	-0.110	0.366***
	(0.347)	(0.471)	(0.328)	(0.0363)
University Connections	-0.421	-0.552	-0.394	0.284***
	(0.295)	(0.401)	(0.280)	(0.0385)
Legal Connections	-1.270**	-1.431*	0.163	0.338***
	(0.571)	(0.774)	(0.533)	(0.0369)
JP	-1.339***	0.212	0.450	0.231***
	(0.471)	(0.559)	(0.340)	(0.0424)
Working Class	-1.305	10.51***	1.524	0.0690
	(1.849)	(2.510)	(1.727)	(0.0561)
Labourer (Family)	2.058	13.79***	-0.104	-0.0365
	(1.896)	(2.584)	(1.768)	(0.0414)
Labourer	0.975	26.07***	0.581	0.224***
	(3.209)	(4.358)	(3.000)	(0.0530)

D.5 Strategic Behaviour: Working Class Constituencies

Table D.5: Baseline Results

	Other × Loss Margin 1906	Working Class × Loss Margin 1906	Loss Margin 1900	Lagged D.V.
Honorary Title	-0.922***	0.456	-0.213	0.344***
	(0.328)	(0.439)	(0.258)	(0.0428)
Military Title	-0.0395	-2.223*	0.734	0.434***
	(0.856)	(1.161)	(0.683)	(0.0468)
Family Hereditary Title	-0.343	-1.245***	-0.603*	0.305***
	(0.373)	(0.473)	(0.322)	(0.0358)
Eton	-0.702	0.0951	-0.185	0.336***
	(0.466)	(0.589)	(0.408)	(0.0363)
Public School	-0.661**	-0.726*	-0.488*	0.316***
	(0.325)	(0.411)	(0.287)	(0.0357)
Fee Paying School	-0.545***	0.455**	0.184*	0.211***
	(0.141)	(0.211)	(0.106)	(0.0516)
Oxbridge	-0.256	-0.901**	-0.449	0.261***
	(0.332)	(0.420)	(0.292)	(0.0378)
University	-0.0696	-0.631*	-0.599***	0.261***
	(0.260)	(0.328)	(0.228)	(0.0383)
Military (Family)	-1.715***	1.337*	0.784	0.295***
	(0.589)	(0.743)	(0.510)	(0.0387)
Military	-0.890**	-0.367	0.108	0.260***
	(0.381)	(0.479)	(0.329)	(0.0376)
Politics (Family)	-1.585***	-0.00513	0.139	0.260***
	(0.425)	(0.536)	(0.369)	(0.0362)
Number Of Clubs	-0.663***	-0.0230	-0.430**	0.243***
	(0.237)	(0.298)	(0.206)	(0.0406)
Brooks Club	-2.844***	1.512	-0.737	0.283***
	(1.050)	(1.322)	(0.908)	(0.0309)
Carlton Club	-0.900***	-0.123	-0.330	0.194***
	(0.254)	(0.322)	(0.224)	(0.0394)
School Connections	-0.583	-0.218	-0.200	0.365***
	(0.376)	(0.475)	(0.330)	(0.0363)
University Connections	-0.103	-1.075***	-0.500*	0.288***
	(0.319)	(0.404)	(0.281)	(0.0384)
Legal Connections	-1.720***	-0.251	0.362	0.340***
	(0.622)	(0.784)	(0.539)	(0.0370)
JP	-1.268***	0.116	0.431	0.231***
	(0.429)	(0.572)	(0.336)	(0.0425)
Working Class	0.580	4.684*	0.618	0.0713
	(2.029)	(2.565)	(1.761)	(0.0567)
Labourer (Family)	1.695	11.73***	-0.207	-0.0168
	(2.069)	(2.613)	(1.795)	(0.0414)
Labourer	2.656	17.54***	-0.580	0.244***
	(3.535)	(4.467)	(3.069)	(0.0538)

E   Results Excluding Controls

This section reproduces the results from the main paper excluding controls: which are the candidate characteristics from 1906 and the loss margin from 1900.

Table E.1: Baseline Results

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.101*	-0.229***	0.179***	0.050
	(0.060)	(0.058)	(0.049)	(0.042)
Adjusted R^2	0.002	0.017	0.015	0.000
N	821	821	821	821

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The 1906 loss margins are the Conservative losses to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

Table E.2: Falsification Test

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
Other Parties × 1906 Loss Margin	0.010	-0.063	0.099*	-0.036
	(0.069)	(0.054)	(0.052)	(0.053)
Conservatives × 1906 Loss Margin	-0.111	-0.166**	0.080	0.086
	(0.096)	(0.076)	(0.073)	(0.074)
Adjusted R^2	0.087	0.183	0.015	0.131
N	1623	1624	1624	1624

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate select in 1910, across different political parties (not just the Conservatives). The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the candidate is from Other Parties (besides the Conservatives) and another interaction with whether candidate is from the Conservatives. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

Table E.3: Don't Lose, Don't Learn?

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	0.378***	0.340**	0.003	-0.343***
	(0.140)	(0.134)	(0.114)	(0.098)
1906 Conservative Defeat $\times$ 1906 Loss Margin	-0.723***	-0.905***	0.293**	0.612***
	(0.173)	(0.165)	(0.140)	(0.121)
1906 Conservative Defeat	-2.956	-1.747	0.081	1.666
	(3.651)	(3.482)	(2.964)	(2.554)
Adjusted R^2	0.021	0.050	0.018	0.029
N	821	821	821	821

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different outcomes in 1906. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin and the 1906 loss margin interacted with whether the Conservatives lost in 1906. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). A dummy variable that takes value of one if the Conservatives lost in that constituency in 1906 is included. Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

Table E.4: Conservative Strongholds

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.049	-0.235**	0.263***	-0.028
	(0.101)	(0.096)	(0.081)	(0.071)
Conservative Stronghold × 1906 Loss Margin	-0.013	0.131	-0.250**	0.119
	(0.133)	(0.127)	(0.106)	(0.093)
Adjusted R^2	0.003	0.025	0.029	0.000
N	821	821	821	821

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different types of constituencies. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the constituency is a Conservative stronghold and the 1906 loss margin interacted with whether the constituency is one of the other constituencies. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

Table E.5: Working-Class Constituencies

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.085	-0.219***	0.214***	0.005
	(0.076)	(0.073)	(0.061)	(0.053)
Working Class × 1906 Loss Margin	-0.042	-0.025	-0.092	0.117
	(0.119)	(0.115)	(0.096)	(0.084)
Adjusted R^2	0.001	0.016	0.015	0.002
N	821	821	821	821

This table reports the estimates of the effects of the 1906 loss margin in a constituency on the type of candidate the Conservatives select in 1910, for different types of constituencies. The dependent variable in Column 1 is the probability that the candidate is a member of the political elite as defined by the Support Vector Machine. In Columns 2, 3, and 4, the dependent variable is the probability that the candidate is a member of the political elite, middle class, and working class, respectively, as defined by the Latent Dirichlet Allocation. The explanatory variables reported are the 1906 loss margin interacted with whether the constituency is a defined as a working-class constituency and the 1906 loss margin interacted with whether the constituency is one of the other constituencies. The 1906 loss margin is the Conservative loss to the winning party, as a percentage of total votes cast in the constituency (negative if the Conservatives won). Standard errors are reported in parentheses and are clustered at the constituency level. ***, **, and * denote statistical significance at 1, 5, and 10 percent level, respectively.

F Classification Robustness

F.1 Support Vector Machine: Different Initial Labels

This section reproduces the Support Vector Machine results in the main paper, but illustrates the robustness of the findings by altering the initial labelling. I instead initially label the unambiguously elite as those whose families have hereditary titles and also attend a public school, while the unambiguous outsiders or working are those whose families are labourers. These results are reported in Table F.1 below.

Table F.1: Baseline Results: Different Initial Labels (SVM)

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-0.401**	0.101	0.383	-0.272	0.00825
	(0.198)	(0.164)	(0.539)	(0.236)	(0.346)
Conservative × 1906 Loss Margin		-0.508*			
		(0.282)			
Conservative Defeat × 1906 Loss Margin			-1.112*		
			(0.602)		
Conservative Stronghold × 1906 Loss Margin				-0.343	
				(0.345)	
Working Class × 1906 Loss Margin					-0.832**
					(0.418)
Adjusted R^2	0.103	0.294	0.107	0.103	0.152
N	468	1038	468	468	238

This table reproduces results from the main paper, with the probabilities of the candidate being part of the political elite defined by a different Support Vector Machine, where the new initial labelling is defined in the above explanation.

F.2 K-Means Clustering

Another potential method of clustering is k-means clustering, which has been used for classification in the economics literature, for example by Crone (2005). For a given number of groups (such as three), k-means clustering minimises the within-cluster variances in the space of candidate characteristics. The three groups that are summarised in Figure F.1, which illustrates a clear political elite, working class, and middle class.

Figure F.1: Characteristics of Candidates in Different K-Mean Clusters
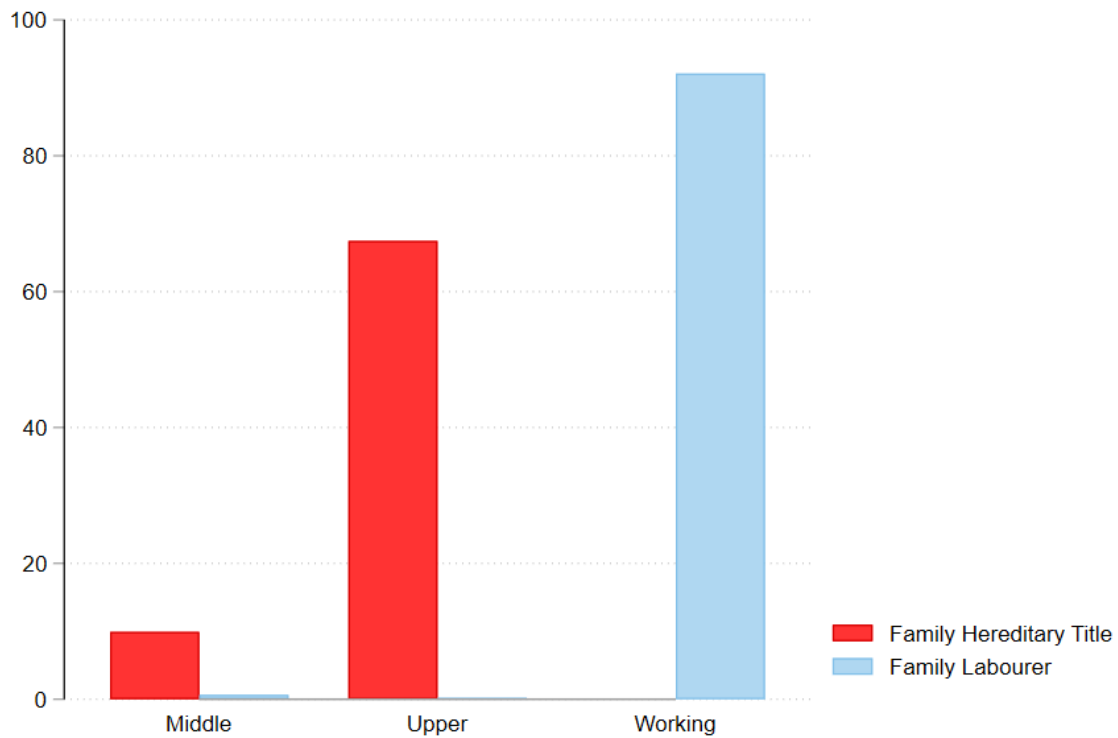
I replicate the results in the main paper, replacing the Latent Dirichlet Allocation variables for the political elite and the working class with indicators from the k-means clustering method, and they are reported in Tables F.2 and F.3.

Table F.2: Baseline Results: K-Means Clustering (Upper Class)

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-1.121***	-0.454	-0.414	-0.926***	-0.301
	(0.231)	(0.325)	(0.637)	(0.256)	(0.440)
Conservative × 1906 Loss Margin		-0.676			
		(0.510)			
Conservative Defeat × 1906 Loss Margin			-1.100		
			(0.718)		
Conservative Stronghold × 1906 Loss Margin				-0.760*	
				(0.430)	
Working Class × 1906 Loss Margin					-1.099**
					(0.498)
Adjusted R^2	0.168	0.230	0.171	0.171	0.172
N	641	1242	641	641	317

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the political elite/upper class as defined by the K-means clustering method.

Table F.3: Baseline Results: K-Means Clustering (Working Class)

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	7.552***	-1.701	-0.991	8.075***	2.618
	(1.566)	(1.045)	(4.286)	(1.738)	(3.960)
Conservative × 1906 Loss Margin		7.059***			
		(1.641)			
Conservative Defeat × 1906 Loss Margin			12.45***		
			(4.815)		
Conservative Stronghold × 1906 Loss Margin				-2.032	
				(2.921)	
Working Class × 1906 Loss Margin					10.72**
					(4.472)
Adjusted R^2	0.0308	0.0286	0.0412	0.0300	0.0593
N	641	1253	641	641	317

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the working class as defined by the K-means clustering method.

F.3 Latent Dirichlet Allocation: Different Number of Groups

I reproduce the results in the main paper using the Latent Dirichlet Allocation for different number of defined groups. The main paper focuses on three groups, and the following reports the results for two, four, five, and six groups. While the results regarding the political elite are not particularly robust, the results regarding the increasing in working-class representation are consistently positive.

Table F.4: Baseline Results: (LDA) 2 Clusters - Political Elite

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-0.337***	0.220***	-0.0296	-0.189	-0.285*
	(0.0879)	(0.0632)	(0.238)	(0.124)	(0.159)
Conservative × 1906 Loss Margin		-0.286***			
		(0.106)			
Conservative Defeat × 1906 Loss Margin			-0.499*		
			(0.267)		
Conservative Stronghold × 1906 Loss Margin				-0.245*	
				(0.146)	
Working Class × 1906 Loss Margin					-0.159
					(0.191)
Adjusted R^2	0.209	0.394	0.216	0.212	0.156
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the political elite/upper class when clustering for two groups using the LDA.

Table F.5: Baseline Results: (LDA) 4 Clusters - Political Elite

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-0.221***	0.121**	0.284	-0.251**	-0.0663
	(0.0821)	(0.0513)	(0.222)	(0.116)	(0.148)
Conservative × 1906 Loss Margin		-0.179**			
		(0.0867)			
Conservative Defeat × 1906 Loss Margin			-0.738***		
			(0.249)		
Conservative Stronghold × 1906 Loss Margin				0.0494	
				(0.137)	
Working Class × 1906 Loss Margin					-0.272
					(0.180)
Adjusted R^2	0.253	0.376	0.268	0.251	0.158
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the political elite/upper class when clustering for four groups using the LDA.

Table F.6: Baseline Results: (LDA) 4 Clusters - Working Class

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	0.0995***	-0.148***	-0.0117	-0.0228	-0.00711
	(0.0377)	(0.0400)	(0.102)	(0.0527)	(0.0754)
Conservative × 1906 Loss Margin		0.0818			
		(0.0677)			
Conservative Defeat × 1906 Loss Margin			0.219*		
			(0.114)		
Conservative Stronghold × 1906 Loss Margin				0.204***	
				(0.0620)	
Working Class × 1906 Loss Margin					0.293***
					(0.0895)
Adjusted R^2	0.0487	0.472	0.0669	0.0684	0.0817
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the working class when clustering for four groups using the LDA.

Table F.7: Baseline Results: (LDA) 5 Clusters - Political Elite

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-0.175**	0.0865*	0.267	-0.121	-0.0314
	(0.0777)	(0.0477)	(0.211)	(0.110)	(0.138)
Conservative × 1906 Loss Margin		-0.128			
		(0.0812)			
Conservative Defeat × 1906 Loss Margin			-0.616***		
			(0.237)		
Conservative Stronghold × 1906 Loss Margin				-0.0887	
				(0.130)	
Working Class × 1906 Loss Margin					-0.302*
					(0.168)
Adjusted R^2	0.216	0.315	0.225	0.215	0.135
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the political elite/upper class when clustering for five groups using the LDA.

Table F.8: Baseline Results: (LDA) 5 Clusters - Working Class

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	0.103***	-0.120***	0.0178	-0.0142	0.0152
	(0.0308)	(0.0333)	(0.0832)	(0.0429)	(0.0641)
Conservative × 1906 Loss Margin		0.0987*			
		(0.0567)			
Conservative Defeat × 1906 Loss Margin			0.175*		
			(0.0930)		
Conservative Stronghold × 1906 Loss Margin				0.195***	
				(0.0505)	
Working Class × 1906 Loss Margin					0.245***
					(0.0765)
Adjusted R²	0.0510	0.486	0.0698	0.0787	0.0913
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the working class when clustering for five groups using the LDA.

Table F.9: Baseline Results: (LDA) 6 Clusters - Political Elite

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	-0.166**	0.0745*	0.222	-0.0809	-0.0272
	(0.0723)	(0.0443)	(0.197)	(0.102)	(0.128)
Conservative × 1906 Loss Margin		-0.129*			
		(0.0755)			
Conservative Defeat × 1906 Loss Margin			-0.534**		
			(0.220)		
Conservative Stronghold × 1906 Loss Margin				-0.141	
				(0.120)	
Working Class × 1906 Loss Margin					-0.281*
					(0.155)
Adjusted R²	0.263	0.343	0.270	0.263	0.190
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the political elite/upper class when clustering for six groups using the LDA.

Table F.10: Baseline Results: (LDA) 6 Clusters - Working Class

	(1)	(2)	(3)	(4)	(5)
1906 Loss Margin	0.143***	-0.0646***	0.0201	0.0402	0.0694
	(0.0263)	(0.0231)	(0.0713)	(0.0367)	(0.0501)
Conservative × 1906 Loss Margin		0.154***			
		(0.0396)			
Conservative Defeat × 1906 Loss Margin			0.188**		
			(0.0797)		
Conservative Stronghold × 1906 Loss Margin				0.171***	
				(0.0432)	
Working Class × 1906 Loss Margin					0.184***
					(0.0605)
Adjusted R^2	0.0995	0.412	0.112	0.127	0.148
N	468	1040	468	468	238

This table reproduces the key results from the main paper, but defines the dependent variable as the probability that the candidate is a member of the working class when clustering for six groups using the LDA.

F.4 Missing Data Assumptions

F.4.1 Missing Biographical Data

This section provides robustness checks concerning missing biographical data on certain candidates. For example, those with missing biographical data may be more likely to come from working-class backgrounds, as they are less likely to appear in the Oxford Dictionary of National Biography if their sole activity in public life was to run for a seat in Parliament. Although it is not immediately clear in what direction the bias due to the missing data goes, I make a range of assumptions about the type of candidates these missing candidates may be, and find that the results are similar to those in the main paper.

As shown in Table F.11, there are a non-negligible percentage of observations where there is incomplete biographical data. Table F.12 replicate the baseline results using the SVM estimates across different assumptions about those with incomplete biographical data, while Tables F.13, F.14, F.15, and F.16 do so for the LDA estimates.

Table F.11: Missing Biographical Data (Percentage of Observations)

	Conservatives	Independents	Labour	Liberals	Average
1900 Election	6.5	66.7	10.0	11.7	8.9
1906 Election	10.9	25	8.2	6.3	8.8
January 1910 Election	14.5	10	7.1	11.0	12.5
December 1910 Election	14.1	40	2.2	11.6	12.5
Average	11.5	26.7	6.3	10	10.7

This table reports the percentage of candidates from each party-election where there is incomplete biographical information.

Table F.12: Missing Biographical Data: SVM Baseline Results

	(1) Missing=0	(2) Missing=10	(3) Missing=20	(4) Missing=30
1906 Loss Margin	-0.368***	-0.341***	-0.314***	-0.288***
	(0.095)	(0.090)	(0.085)	(0.081)
1900 Loss Margin	-0.102	-0.087	-0.073	-0.059
	(0.104)	(0.098)	(0.093)	(0.088)
Lagged Dependent Variable	0.204***	0.212***	0.219***	0.225***
	(0.040)	(0.039)	(0.039)	(0.038)
Adjusted R^2	0.080	0.081	0.082	0.082
N	641	641	641	641

This table replicates the baseline results assuming different percentage probabilities that the candidates who have incomplete biographical data are from the political elite, where those who have complete biographical data have their probabilities estimated by the SVM. For example, in Column 1, the dependent variable is assumed to be 0 where there is incomplete biographical data.

Table F.13: Missing Biographical Data: SVM Baseline Results - Scenario 1

	(1) Elite (LDA)	(2) Middle (LDA)	(3) Working (LDA)
1906 Loss Margin	-0.345***	0.020	0.346***
	(0.077)	(0.066)	(0.084)
1900 Loss Margin	-0.037	0.015	0.061
	(0.083)	(0.072)	(0.091)
Lagged Dependent Variable	0.329***	0.253***	0.207***
	(0.037)	(0.041)	(0.041)
Adjusted R^2	0.155	0.053	0.078
N	641	641	641

This table replicates the baseline results assuming Scenario 1: that all of those who have incomplete biographical data are 100 percent likely to be in the working class, and 0 percent likely to be in the political elite or the middle class. Those with complete biographical have their probabilities calculated by the LDA, as in the main paper.

Table F.14: Missing Biographical Data: SVM Baseline Results - Scenario 2

	(1) Elite (LDA)	(2) Middle (LDA)	(3) Working (LDA)
1906 Loss Margin	-0.345***	0.045	0.317***
	(0.077)	(0.063)	(0.077)
1900 Loss Margin	-0.037	0.025	0.044
	(0.083)	(0.068)	(0.084)
Lagged Dependent Variable	0.329***	0.269***	0.224***
	(0.037)	(0.040)	(0.040)
Adjusted R^2	0.155	0.063	0.084
N	641	641	641

This table replicates the baseline results assuming Scenario 2: that all of those who have incomplete biographical data are 90 percent likely to be in the working class, 10 percent likely to be in the middle class, and 0 percent likely to be in the political elite. Those with complete biographical have their probabilities calculated by the LDA, as in the main paper.

Table F.15: Missing Biographical Data: SVM Baseline Results - Scenario 3

	(1) Elite (LDA)	(2) Middle (LDA)	(3) Working (LDA)
1906 Loss Margin	-0.319***	0.045	0.289***
	(0.073)	(0.063)	(0.070)
1900 Loss Margin	-0.025	0.025	0.027
	(0.079)	(0.068)	(0.076)
Lagged Dependent Variable	0.346***	0.269***	0.243***
	(0.036)	(0.040)	(0.039)
Adjusted R^2	0.164	0.063	0.092
N	641	641	641

This table replicates the baseline results assuming Scenario 3: that all of those who have incomplete biographical data are 80 percent likely to be in the working class, 10 percent likely to be in the middle class, and 10 percent likely to be in the political elite. Those with complete biographical have their probabilities calculated by the LDA, as in the main paper.

Table F.16: Missing Biographical Data: SVM Baseline Results - Scenario 4

	(1) Elite (LDA)	(2) Middle (LDA)	(3) Working (LDA)
1906 Loss Margin	-0.319***	0.069	0.262***
	(0.073)	(0.060)	(0.064)
1900 Loss Margin	-0.025	0.035	0.012
	(0.079)	(0.065)	(0.070)
Lagged Dependent Variable	0.346***	0.281***	0.264***
	(0.036)	(0.040)	(0.039)
Adjusted R^2	0.164	0.073	0.101
N	641	641	641

This table replicates the baseline results assuming Scenario 4: that all of those who have incomplete biographical data are 70 percent likely to be in the working class, 20 percent likely to be in the middle class, and 10 percent likely to be in the political elite. Those with complete biographical have their probabilities calculated by the LDA, as in the main paper.

F.4.2 Uncontested Elections

The baseline analysis in the main paper excludes observations where the Conservative candidate in either the 1900 or the 1906 election ran uncontested, as the 1900 Loss Margin and the 1906 Loss Margin are not unambiguously defined. As a robustness check, I reproduce the baseline results in Table 2, but also code 1900 and 1906 Loss Margins as -100 if the Conservative candidate ran unopposed (a 100 per cent majority), and 100 if no Conservative candidate ran. The results are reported in Table F.17. Although the magnitudes are a little smaller than in the main analysis, they still provide significant evidence of an association between a greater loss margin and switching candidates from the political elite to the working class.

Table F.17: Baseline Results: Including Uncontested Elections

	(1) Elite (SVM)	(2) Elite (LDA)	(3) Middle (LDA)	(4) Working (LDA)
1906 Loss Margin	-0.102*	-0.110**	0.026	0.081*
	(0.062)	(0.056)	(0.048)	(0.041)
1900 Loss Margin	-0.022	-0.057**	0.061***	-0.001
	(0.028)	(0.026)	(0.022)	(0.019)
Lagged Dependent Variable	0.355***	0.488***	0.465***	0.439***
	(0.034)	(0.032)	(0.035)	(0.031)
Adjusted R^2	0.133	0.260	0.211	0.208
N	744	744	744	744

This table replicates the baseline results in the main paper, but assumes that the loss margin in an election where the Conservative Party ran unopposed is -100, and 100 if the Conservative Party did not run at all.