UC Berkeley
CEGA Working Papers

Title
Who Becomes a Politican?

Permalink
https://escholarship.org/uc/item/06v8f1rp

Authors
Dal Bo, Ernesto
Finan, Frederico
Folke, Olle
et al.

Publication Date
2016-10-21

eScholarship.org Powered by the California Digital Library
University of California

Series Name: WPS
Paper No.: 063
Issue Date: 21 October 2016

Who Becomes a Politician?

Ernesto Dal Bo, Frederico Finan, Olle Folke,
Torsten Persson, and Johanna Rickne

Working Paper Series
Center for Effective Global Action
University of California

This paper is posted at the eScholarship Repository, University of California. http://escholarship.org/uc/cega_wps
Copyright © 2012 by the author(s).

The CEGA Working Paper Series showcases ongoing and completed research by faculty affiliates of the Center. CEGA Working Papers employ rigorous evaluation techniques to measure the impact of large-scale social and economic development programs, and are intended to encourage discussion and feedback from the global development community.

Recommended Citation:

Dal Bo, Ernesto; Finan, Frederico; Folke, Olle; Persson, Torsten; Rickne, Johanna. Who Becomes a Politician? CEGA Working Paper Series No. WPS-063. Center for Effective Global Action. University of California, Berkeley.

Who Becomes a Politician? *

Ernesto Dal Bó, Frederico Finan, Olle Folke,
Torsten Persson, and Johanna Rickne

August 2016

Abstract

Can a democracy attract competent leaders, while attaining broad representation? Economic models suggest that free-riding incentives and lower opportunity costs give the less competent a comparative advantage at entering political life. Also, if elites have more human capital, selecting on competence may lead to uneven representation. We examine patterns of political selection among the universe of municipal politicians in Sweden using extraordinarily rich data on competence traits and social background for the entire population. We document four new facts: First, politicians are on average significantly smarter and better leaders than the population they represent. Second, the representation of social background, whether measured by intergenerational earnings or social class, is remarkably even. Third, there is at best a weak tradeoff in selection between competence and representation. Fourth, both material and intrinsic motives matter in selection, as does screening by political parties.

*We are grateful to seminar and conference participants at Berkeley, Bocconi, CAF, CIFAR, Columbia, IIES, EIEF, LACEA-Political Economy, LBS, LSE, NBER, PECO, and Stanford for helpful comments. We thank Daniele Paserman, Martín Rossi and Juan Vargas for their discussions of the paper, and gratefully acknowledge the Söderberg Foundation and the Swedish Research Council for financial support. Johannes Hermle provided excellent research assistance.

1 Introduction

The identity of political leaders affects which policies get selected, how well they are implemented, and who benefits from them.[1] While this is intuitive for autocracies where rulers face few constraints, it is also true for representative democracies, because policy platforms do not constitute complete enforceable contracts. Most voters would therefore like to elect highly competent policymakers for choosing and implementing policies to attain a given objective (throughout the paper, we use the terms "able" and "competent" interchangeably). As a collective, voters may also want to elect policymakers who represent diverse interests, so that government will pursue broad objectives.

Whether representative democracy can deliver both high-ability leaders and broad representation is unclear. Economic models of politics suggest that the less able have a comparative advantage at entering public life due to lower opportunity costs (Caselli and Morelli 2004) and that this effect may be compounded by free-riding incentives (Messner and Polborn (2004), see also Olson (1965)). In addition, if ability is distributed unequally, selecting on competence may make it harder to ensure broad representation. Related to this, a number of scholars have argued that the electoral system shapes the tradeoff between accountability – a driver of selection – and representation.[2]

To better understand political selection, and the potential tradeoffs it entails, we need to thoroughly describe selection patterns and analyze their determinants. Unfortunately, insufficient data has made it difficult to carry out these tasks.

Three data limitations First, any study of political selection should account for two stages, namely candidate entry and screening by voters and/or parties. To study candidate screening one needs information on both elected and non-elected politicians. While information on the former is sometimes available, information on the latter is remarkably sparse. A few studies have tried to tackle this limitation to advance our knowledge of candidate selection.[3] Unfortunately, this literature

[1] See for example Osborne and Slivinski (1996), Besley and Coate (1997), Pande (2003), Chattopadhyay and Duflo (2004), Jones and Olken (2005), Washington (2008), Besley, Montalvo, and Reynal-Querol (2011), Meyersson (2014).

[2] A common idea is that plurality rule fosters better accountability, while proportional representation fosters better representation (Myerson 1993; Persson and Tabellini 2003; Powell 2000; Taagepera and Shugart 2000).

[3] Examples include: Besley, Pande, and Rao (2005), Ferraz and Finan (2009), Galasso and Nannicini (2011), Beath, Christia, Egorov, and Enikolopov (2014), Jia, Kudamatsu, and Seim (2015), and Tillmann (2014).

does not address candidate entry, which requires information on those who never attempted to enter into politics.

Second, much of the relevant theory emphasizes the quality of political selection. But how to measure ability? Absent direct data on the underlying intelligence or personality traits of politicians, the existing empirical literature has relied on proxies such as education or pre-office income (Dal Bó, Dal Bó, and Snyder 2009; Ferraz and Finan 2009; Merlo, Galasso, Landi, and Mattozzi 2010; Galasso and Nannicini 2011; Besley and Reynal-Querol 2011; Gagliarducci and Paserman 2012). While correlated with ability, these proxies also likely reflect luck or social class.

Third, representation is hard to measure. Previous work has relied on proxies like occupation. However, occupation is also coarse: many politicians may be lawyers, but if values and loyalties depend on social background one would like to know if they were brought up as elite or working class. This requires intergenerational information that is difficult to obtain.

In sum, we know of no research that has been able to analyze both stages of the selection process, while using a comprehensive bundle of traits that accurately reflect the ability and representativeness of politicians.

Sweden as a test bed Our study overcomes these limitations by using fine-grained administrative data from Sweden. This focus deserves some justification beyond data availability. Clearly, conclusions based on evidence from a single country cannot be extrapolated to the rest of the world, although we take a limited comparative step by studying different Swedish jurisdictions. But Sweden is of interest for both its commonalities and its differences with other countries. Swedish politics is based on an electoral system with proportional representation, the most common type of democracy in the world. Sweden is also a quintessential advanced democracy. It has remained stable and fully democratic with a top score on the 20-step Polity-IV scale since 1917, a century during which other states gradually moved from autocracy towards democracy. Sweden thus exemplifies an institutional destination many countries around the world may be approaching. When debating the value of democracy it is natural to ask if a democratic transition can improve the ability and representativeness of leaders. If Sweden displayed incompetent and unrepresentative leaders, advocates of democratization may have to resort to different arguments.

Our data The Swedish data allows us to undertake the most thorough description to date of the basic patterns of a country's political selection. A key advantage is the availability of rich measures of individual ability and social background. The ability measures include evaluations of IQ and leadership potential done by the military on

the universe of the male population, and an estimate of residual ability for the full population stemming from a fully saturated Mincer regression developed by Besley, Folke, Persson, and Rickne (2013). Our social-background data entails reliable intergenerational information, namely social class as reflected in parental incomes and occupations (a traditional approach in sociology).

We study politicians in municipalities – which provide the bulk of public services – to take advantage of large numbers and within-country variation. We analyze these data at three different levels that make for the three main sections of the paper.

Aggregate level At the highest level, we uncover new statistics on political selection for Sweden as a whole. Are politicians positively or negatively selected in terms of ability? Standard models of occupational choice would predict adverse selection: able people have higher private incomes and more promising careers and hence face higher opportunity costs to enter public life. Nevertheless, we document strong positive selection along all ability measures. Selection monotonically improves from those nominated but not elected – via rank-and-file elected council members – to mayors. Mayors are almost as positively selected as national legislators and members of elite occupations in the private and public sectors.

We then examine the social origin of politicians. While politicians themselves are disproportionately high-earning, their parents' social class and earnings are distributed almost identically to those of the population. In other words, politics is not reserved for the scions of elite families but tracks the entire social structure. This partly reflects different political parties representing different segments of the income distribution.

The findings on representation help us understand the drivers of selection on ability. A priori, positive selection could reflect either meritocracy that works independently of social class, or elitism that works via privileged access to political power and to human capital. If elitism is the main driver of access to political office, family background should strongly predict selection and, conditional on that background, individual traits should not. Our representation findings indicate that family background is not an important driver. Moreover, when comparing the abilities of politicians and their siblings, we find differences nearly as large as those separating politicians from the general population. Conditional on family background, individual traits thus matter greatly for selection.

Municipal level In the second part, we examine municipal heterogeneity. Here, we test whether municipalities that select more strongly on ability do so at the cost of narrower representation. We document substantial variation across municipalities

in each dimension of selection, but find no evidence of an acute tradeoff. Upon closer inspection, the flat tradeoff reflects stronger selection on ability among lower socioeconomic strata.

Individual level Finally we consider individual decisions about entry into politics, and the role political parties may play in giving them access to political office. A stylized formal model – which shares some features with the aforementioned economic-selection models – guides our thinking on drivers of self-selection. It predicts (i) higher mayoral wages to attract more competent individuals not only to the mayor position, but also to engage in politics. The model also predicts the entry motives into politics to be enhanced by (ii) intrinsic motives for public service, and (iii) lower returns to seniority in one's private career. The correlations we find in the data are consistent with these three predictions.

Our analysis also indicates that, given self selection, the positive selection on ability likely reflects screening by political parties, allowing the more able in and promoting them to positions of influence. This finding matches qualitative work in political science. It also matches two facts in our data, namely that ability goes up by list rank within parties, and that ability for each list rank goes up by levels of political competition.

Organization of the paper In the next section, we offer background information on the Swedish political system. Section 3 describes our data and their sources. In Section 4, we present the aggregate results on political selection of ability and social origin. Section 5 explores the two dimensions of selection across municipalities, and examines the prospective ability-representation tradeoff. In Section 6, we analyze individual self-selection, both theoretically and empirically, and discuss the role of parties as screening devices. We conclude in Section 7.

2 Background

Sweden's electoral system Sweden has three administrative levels of government. Every four years (three years prior to 1994), elections are run for 290 municipalities, 20 counties, and the nation. All elections take place on the same day with a turnout between 80 and 90 percent. In each election, citizens cast a separate party ballot, a ranked list with a large number of candidates. This system elects a total of 349 parliamentarians, 1,100 county-council members, and 13,000 municipal-council members. Our paper is focused on the latter group.

In Sweden's proportional-representation (PR) system, seat shares in the municipal councils closely trace the vote shares of political parties. Until 1998, seats for each party were allocated from the top of the ballot, i.e., in typical fashion under a closed-list system. Since 1998, voters can also cast an optional preference vote for one candidate. But this reform has only allowed a handful of politicians from lower ranks to bypass the party's list order and win a seat.[4]

Based on the distribution of council seats, a ruling majority is formed. These majorities often form within the left bloc (Social Democrats, Left Party, and Green Party) or the center-right bloc (Conservatives, Center Party, Liberal Party, and Christian Democrats). Occasionally, the largest local party wins more than half of the seats and rules on its own. In our time period, two anti-immigration parties (New Democracy, in the 90s, and Sweden Democrats, in the 00s) have also been represented, but these parties are rarely part of governing majorities. Local parties (running only in one municipality) also exist, but usually hold less than 5 percent of the seats.

Municipal governance The council is the only directly elected body in the municipality. It has a board – the local analog of the national cabinet – the members of which are appointed by the governing majority to mirror the seat distribution. The largest party in the ruling majority selects the chair: henceforth, the "Mayor." The political opposition usually appoints an executive as well, the "Vice Mayor." Mayors play an important role in municipal governance.

The mayor is an important office, since municipalities play a crucial role as service providers in the economy: they carry out expenditures of about 25 percent of Sweden's GDP and employ 20 percent of the country's workforce. Municipal politicians are thus responsible for the areas of K-12 education, child care, elderly care, and local infrastructure, and finance these commitments through a local income tax at around 20 percent of income.

Ruling over the Swedish welfare state used to be a Social Democratic prerogative. But in the past few decades, political competition has grown substantially more intense. This can be illustrated by the changing proportion of left-bloc governments over time: it increased from 31 percent in 1991 to 73 percent in 1998, only to fall back down to 59 percent in 2002 and 41 percent in 2006.

[4]This reflects voter "abstention" from the optional vote, a concentration of votes for candidates at the top of the ballot, and high thresholds. See Folke, Persson, and Rickne (2015) for a thorough analysis of the preference-vote system and its consequences.

Running for office and political careers Sweden has a strong normative tradition of citizen-politicians, in Swedish called the "leisure politician". The people's representatives should step into the political gathering directly from the farm, the factory, or the office to represent his or her peers in the management of their joint resources. The political appointment should be handled in ones' spare time, contrasting the idea of "professional" politics where positions require specialized, elite knowledge. Reflecting the ideal of the citizen politicians, almost all political offices are honorary. With the exception of the 349 full-time parliamentarians, the only other full time political positions are the top executives at the local levels. In larger regions, there also exists vice chairs (most importantly the vice mayor) who is paid a part-time salary. Previous research has shown that there are no monetary gains for winning a seat in a municipal assembly (Lundqvist 2013). On the other hand, becoming the mayor is associated with higher status and substantial power. The wage of the mayor is in the top percentile of the Swedish earnings distribution, and - curiously - this wage is determined in the municipal council.

Becoming a politician may seem easy enough, but the idea of highly motivated persons who actively join parties, show off their skills and climb the ranks is not fitting for the Swedish recruitment procedure (Nielsen 2001) , see also Teorell (2000). This individualistic view does not rhyme with a party-centered policy making process where politicians toe the party line and are dependent on elite support to climb the career ladder (Öhberg 2010). Swedes are more likely to wait for an external push from friends, family or party recruiters to enter politics, rather than volunteering (Karlsson 2001). Interviews with local party recruiters show that they selectively seek out individuals that they deem suitable for office (Soininen and Etzler 2006). Fredriksson (2001) found that one in four local politicians had been directly targeted by the party after some noticeable action in the locality, for example starting a political organization or advocating forcibly for an issue.

Monetary costs to running for municipal office are minimal. All politicians run as part of a party ballot, and campaign finance is mostly paid by tax money and channeled to parties rather than candidates. Individual campaigns for preference votes have modest costs, with the vast majority spending less than 600 USD at the municipal level. Even these small costs are usually paid by the party or by outside donors, rather than by the individual herself (SOU:68 2007).

Even if recruitment happens at someone else's initiative, unwilling candidates can - of course - not be forced into office. Several interview and survey studies have explored motivations for going into politics. In written surveys, politicians overwhelmingly attribute their decision to a wish to influence social or economic outcomes in accordance with some ideological goal(s) (Karlsson 2001). They also

cite appreciation for the social context of the political party, and the ambition to learn more about government. A common understanding of political recruitment is that citizens, and later party members, reluctantly say yes when they are asked to participate or take on an appointment, but altruistically do so in order to work for the betterment of society (Erlingsson and hrvall 2012).

A complementary viewpoint is offered by Dahl (2011)who interviews local politicians to find that the motivation for a political career grows among local politicians once they have taken the initial step into a political appointment. At the early stage of party membership, people are attracted by the social environment and by interests in particular policy areas such as education or child care. Over time as party members, and particularly when gaining an elected or appointed office (which happens after 7-8 years on average, see (Karlsson 2001)), ambitions for higher office intensify. Having decided to spend their free time on politics, most persons with an appointment in Dahl's study had considered (or dreamed) about the opportunity to turn their hobby into a job. When asked about others' ambitions, responses were even more frank. Most respondents believed that others joined the party in order to gain influence via increasingly powerful appointments. Respondents also referred to monetary rewards, with one recruiter claiming that "many have ulterior motives and become active for their private benefit. They do it to be seen or to get money, today or in the future" (Dahl (2011), p. 186).

Nevertheless, the politicians were aware that their promotion chances hinged on factors "outside of their own control", most importantly the support of their fellow party-members and the party's electoral performance ((Dahl 2011), p 184). This quote highlights the particular role of intra-party support as a basic building block for a political career in Sweden's party-centered system (Niklasson 2005). The key event in the political cycle, and where this support is put to use, is the procedure that decides who should be on the electoral ballot and, even more importantly, in which slot. The list's top ranked individual leads the party's formulation of policy proposals and acts as its political spokesperson, and as noted above, it is standard procedure is appoint him or her to the top executive position available to the party. The top person is followed on the list by the remaining candidates in an order that ranks their political influence. Climbing to a higher list rank is, essentially, an upward career move, because the top portion of the list is elected, and the top portions of the elected politicians are given executive appointments.

Municipal parties compose their own electoral ballots without interference of the central party organization. Within the municipality, members are organized into clubs based on either geographical areas or party sub-organizations. Large parties typically have clubs for the youth league, women, university students, and (in left-

wing parties) trade-unionists. In the two left-wing parties (Social Democrats and Left Party), clubs nominate members to a committee that aggregates the nominations into a proposed list. In the other parties, the nomination committee organizes test elections, in which all party members can vote. On the face of it, such an internal primary may seem open to turnover within the list, but in reality both nomination procedures strongly favor the status quo. Dominant factions within the party control the aggregation of nominations, both through the clubs and the committee which makes up the final list. In the internal primaries, votes are mobilized by strong actors and networks, which organize around candidates. These networks are important, since only a minor fraction of the members use their right to vote.

Voter preferences and selection Swedish citizens say that they value a competent and socially representative political class. Asked about what drives their party choice, politician competence has remained a top-three reason in the past decades, rising to the same importance as ideology, and surpassing issue voting (Statistics Sweden 2010). Asked about what social dimensions should be reflected in influential positions, voters rank gender the highest, closely followed by age, social group, and geographic area (Djerf-Pierre and Niklasson 2010).

Representation of social classes partly occurs through the party system. Traditionally, left-bloc parties represent blue-collar workers, while center-right parties represent white-collar workers. In a recent survey of newly elected municipal and county politicians, 48 percent of Social Democrats classified themselves as working class and 42 percent as middle class. Among Conservatives, 5 percent saw themselves as working class, 42 percent as middle class, and the rest as upper class. Outside these two large parties, the Center party has traditionally represented farmers. Within parties, social representation is advanced via strategies to increase the representation of women, foreign-born, and the young (Freidenvall 2006).

3 Data

To characterize political selection, we assemble the (to our knowledge) most detailed and comprehensive data set to date. In this section, we briefly summarize our sources, key variables, and sample definitions.

Sources Our empirical analysis is based on individual-level data from various sources. Our first dataset contains all elected and non-elected individual candidates that ran for political office during the period of 1982-2010, about 200,000 unique

individuals. Prior to each election, each political party must report its ordered list, together with the personal identification code for each listed politician. These lists are kept by Statistics Sweden and, in some cases, regional electoral authorities. After each election, another record is created with a complete list of all elected politicians from each party. Altogether, our sample has about 50,000 elected individuals.

We link these data on politicians to several administrative registers held by Statistics Sweden. These databases hold annual data for the whole Swedish population, aged 16 and above. For most variables, our data holds annual records from 1979 to 2012 for the entire population, altogether approximately 14 million unique men and women. These data contain detailed demographic and background information (e.g., age, sex, education level, occupation, etc.), as well as earnings. With this information in hand, we can precisely characterize how the personal attributes of politicians relate to those in the entire population.

We use data from the Multigenerational Register to measure intergenerational linkages, which allows us to identify the siblings and parents of a politician. We use only biological parents, and as the data begins in 1979, we face a natural truncation. Nevertheless, for politicians elected in 2010, we observe at least one parent for 91 percent of the sample.

Various types of earnings are available from the Swedish Tax Authority on an annual basis for the entire population. We also have universal annual information about individual sector of employment for the whole period. However, information of occupation is only recorded on a yearly basis from 2003. To bridge this gap for earlier periods, we further complement the occupation data with information from Population Censuses (available every five years).

Our final individual-level dataset comes from the Swedish Defense Recruitment Agency. For a subset of cohorts, scores from the military enlistment procedure offer statistics on the mental capacities of Swedish men. Although the mandatory draft was instituted in 1901, full records of scores are only kept for cohorts born 1951 and onward. For quality reasons, we also truncate the data for men born after 1980. For these 30 cohorts, enlistment rates were around 90 percent.

Electoral results are linked to our dataset at the municipal level by using records kept by the Swedish Electoral Agency. These give us the vote shares for every party in every election. Data on the party that appointed the mayor was obtained from Kfakta, a database collected by Leif Johansson (Department of Political Science, University of Lund)

The enlistment procedure The enlistment process for military service spans two days and evaluates a person's medical status, physical fitness, and cognitive abilities.

About 75 percent of the people in our sample who took the test did so at 18 years of age. Most of the remaining 25 percent were 19 years old (less than 0.5 percent were below 18 or above 19). Enlistment generally occurs in the year a person graduates from high school, so test scores are not influenced by university training. Because these tests tended to be high-stakes – better results gave more rewarding military placements – the data quality is considered high. Takers were not informed of their specific scores.

We use two scores from the enlistment procedure – the cognitive score and the leadership score. Each of these measures is standardized to a (stanine) scale from 1 to 9.

Cognitive score Cognitive ability is scored based on written test with four subsections. The subtests assess the person's ability in the domains of problem solving, induction capacity, numerical ability, verbal ability, spatial ability, and technological comprehension (Ståhlberg-Carlstedt and Sköld 1981). Army expert Berit Carlstedt (2000) argues that the Swedish enlistment test is a good measure of general intelligence. This differentiates the Swedish test from other tests, such as the US Armed Forces Qualification Test, which focuses more on "crystallized" intelligence, i.e., teachable skills. We can thus think of the cognitive score as the result of an IQ test. The scale is standardized such that a score of 5 is reserved for the middle 20 percentiles of the population taking the test, while the scores of 6, 7, and 8, are given to the next 17, 12, and 7 percentiles, and the top score of 9 to the uppermost 4 percentiles (scoring below 5 is symmetric).[5]

Leadership score Individuals who score a 5 or higher on the cognitive test go through a more in-depth evaluation for possible placement into a military leadership position. Trained psychologists administer a semi-structured interview to determine the 1-9 leadership scores. Before the interview, the psychologist receives information about the conscript's cognitive test results, physical endurance, muscular strength, grades from school, and the answers to 70-80 questions about friends, family, and hobbies, etc. The exact manual used in the interviews is classified information, but the test is known to evaluate a conscript's personality in civilian life, and his ability to handle military functions. Specifically, the leadership score summarizes four personality traits: social maturity, psychological energy, intensity, and emotional stability. These are closely related to the well-known big-5 personality traits (extraversion,

[5]In terms of standard deviations, scores translate as follows 1: below -1.75, 2: -1.75 to -1.25, 3: -1.25 to -0.75, 4: -0.75 to -0.25, 5: -0.25 to 0.25, 6: 0.25 to 0.75, 7: 0.75 to 1.25, 8: 1.25 to 1.75, and 9: above 1.75.

consciousness, openness, conscientiousness, and neuroticism) – see the Appendix, Table A.1.

Education Educational attainment of all individuals is reported by Swedish schools and universities, and records are kept by Statistics Sweden. For people immigrating into Sweden later in life, information on schooling is collected through surveys and education levels are categorized into a Swedish standard. These standard categories can easily be translated into years of education.

Income We use a measure of annual disposable income, which is constructed from individual tax records (Sweden does not have joint family taxation). This variable includes all income sources and all government transfers. Thus, it includes wages and in-kind benefits from jobs, pensions, transfers and subsidies, business income, capital income, sickness and parental-leave benefits, etc.

Residual ability To gauge relative earnings power, which we refer to as residual ability, we follow the approach in Besley, Folke, Persson, and Rickne (2013). These authors use residuals from a fully saturated Mincer equation, defined over a large set of socioeconomic characteristics. For computational reasons, this equation is estimated year by year. Specifically, we estimate the equation:

$$y_{i,m,t} = f(age_{i,t}, educ_{i,t}, empl_{i,t}) + \alpha_m + \varepsilon_{i,m,t} \ , \tag{1}$$

where the dependent variable $y_{i,m,t}$ is the disposable income for person i in municipality m in year t. Among the independent variables, $age_{i,t}$ represents a set of age indicators (over 5-year intervals), $educ_{i,t}$ is a binary indicator equal to zero when the individual has less than tertiary education and equal to one otherwise, and $empl_{i,t}$ denotes a set of indicators for 15 activity codes.[6] The function f represents a full

[6]These are the same as the European NACE code and international ICIC code, namely: "Agriculture, hunting and forestry", "Fishing", "Mining and quarrying", "Manufacturing", "Electricity, gas and water supply", "Construction", "Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods", "Hotels and restaurant", "Transport, storage and communication", "Financial intermediation", "Real estate, renting and business activities", "Public administration and defense; compulsory social security", "Education", "Health and social work" and "Other community, social and personal service activities". Two categories, "Activities of households" and "Extra-territorial organization and bodies" have less than 30 individual-year observations in them. Because of this, we add the former to "Other community, social and personal service activities", and the latter to "Public administration and defense; compulsory social security".

set of age-education-employment interactions. The regression also includes municipality fixed effects α_m to capture systematic income differences across regions, or between urban and rural areas. All in all, this flexible specification allows for different earnings-tenure profiles across sectors and education levels. For each individual, we compute the residual $\varepsilon_{i,m,t}$ for each available year in the panel, and then average across years. The so-computed "individual fixed effect" is our residual ability measure.

To minimize measurement error and endogeneity, we drop observations for full-time politicians, both in office and after exiting office. We also do two sample divisions when estimating equation (1), namely by gender and retirement status (age over 65 or not) in order not to confound the competence measure with the substantial differences in labor-market behavior across these groups.[7]

Does residual ability measure a competence for politics rather than an idiosyncratic ability to generate market income? Besley, Folke, Persson, and Rickne (2013) address that question and show that residual ability is indeed correlated with different measures of political success, of cognitive and leadership ability, and of policy success.

Summary statistics Table 1 displays pair-wise correlations between our main competence measures for the Swedish male population. All measures are positively correlated. As expected, years of schooling and the cognitive score are the most strongly correlated with a correlation coefficient of 0.51. Schooling and leadership skills are less strongly related at 0.30. Residual ability, which already controls for education, has the weakest association with the other measures. In sum, the four measures in the table appear to capture different dimensions of ability. While it is an open question which is the best measure of competence, these simple correlations highlight the hazards of relying solely on years of schooling.

[Table 1 about here]

Table 2 reports means of the ability variables for politicians and the entire population (subject to availability) in the year of 2011. For municipalities, we distinguish between the nominated (but non-elected), elected, and mayors. For comparison's

[7] As for gender, more than 30% of women who work do so part time, in contrast to less than 10% of the men. Women also take on the bulk of parental leave and care activities that raise the gender pay gap when couples have children. As for retirement, there are plenty of senior politicians. Mincer equations of retired and working people differ as retirees do not have a current employment sector. We compute the income residuals of retirees based on the main sector of employment in their working-life.

sake, we also include information on national legislators (MPs). Compared to the population, Swedish politicians tend to underrepresent women and the foreign born. Of mayors, less than 30 percent are women and less than 3 percent foreign born. We do not address these forms of under-representation any further in this paper. The most striking point of the data is a main topic in our analysis to follow. Swedish politicians are positively selected, based on all four ability measures. The progression of mean cognitive and leadership scores from nominated to elected to mayor suggests increasing rates of positive selection. National legislators have similar cognitive and leadership scores as mayors, but hold a full extra year of education.

[Table 2 about here]

4 Selection in the Aggregate

We begin by considering how Swedish politicians are selected on average, with regard to competence as well as social background. The analysis presented in this section thus relies on the universe of municipal politicians.

4.1 Positive or Adverse Selection?

Politicians vs. the population Our first main contribution is to compare the ability characteristics of politicians to those of the general population. We study the four different ability dimensions introduced in Section 3. To repeat, two measures – education and residual ability – are available for the entire population, while the other two – cognitive and leadership scores – are available only for males.

Leadership and cognitive scores The top-left panel in Figure 1 shows overlapped histograms for the leadership scores of the general (male) population and three categories of municipal politicians – nominated but not elected, elected, and mayors. A clear pattern of positive selection emerges. The distribution of leadership scores of the nominated looks quite close to that of the population but with a slight shift to the right – cognitive scores above the population mean of 5.2 are more highly represented among nominated politicians than in the general population.[8] For elected politicians, the shift to the right is stronger, and even more so for

[8] All differences across groups reported in this subsection are strongly statistically significant, with p-values below 0.001. A similar pattern holds in other sections of the paper unless noted otherwise.

mayors. Mayors have more than a full additional point (out of nine) – almost 3/4 of a standard deviation – higher leadership scores than the population.

[Figure 1 about here]

The top-right panel shows a similar graph for the cognitive score. Politicians score higher than the average Swede, more strongly so when elected to office, and particularly so when selected for top-municipal office. Mayors score an additional 1.3 points higher than the average person (who took the test), about 2/3 of a standard deviation in the population.

Residual ability and education The bottom-right panel of Figure 1 displays the distributions of residual ability. The nominated display a small shift to the right – their mean residual ability is higher by 1/8 of a standard deviation in the population. The elected show a clearer shift, with a 1/2 standard-deviation difference. Mayors have the highest residual ability levels, with a mean almost a full standard deviation above the population mean.

The evidence on residual ability is important for two reasons. First, it includes both females and males. Second, strong positive selection on intelligence and leadership alone might just reflect a lower opportunity cost for those who become politicians. But the opposite seems to be true, as politicians also have higher residual ability (a measure driven by earnings) as well as actual pre-office earnings – see Figure 3 below.

The distribution of education attainment over seven levels (in the bottom-left panel of Figure 1) shows a similar pattern, with politicians underrepresented at the bottom two levels, and over-represented at higher levels. As reported in Table 2, the nominated have one more year of education than the average Swede, whereas elected politicians and mayors have 1.3 years extra. In contrast to the other ability measures, this politician distribution does not first-order stochastically dominate the population distribution: politicians are sightly underrepresented at the highest level, which includes PhDs and makes up a minimal share of the population. In other words, politicians are positively selected in terms of education levels and years, but doctoral–type degrees are less prevalent among politicians. In the remainder of the paper, we focus on the non-education measures of competence.

The key takeaway from these graphs is a strong pattern of positive selection in Swedish politics, which gets more positive at higher positions in the political hierarchy. This flies in the face of the fact that more competent people have a higher opportunity cost of entering public life.

Politicians vs. other high-status professions To gain an additional perspective on selection, we also compare politicians to members of Swedish elite occupations, known for attracting talented people. Table 3 shows our competence measures as well as earnings for mayors, politicians elected to municipal council, CEOs, lawyers, medical doctors, and academic social scientists. The positive selection among CEOs increases with company size. Elected politicians have cognitive and leadership scores similar to CEOs with 10-25 employees, a group which is also comparable in size. Mayors have exactly the same scores as CEOs in companies with 25-250 employees, even though mayors earn substantially less. Lawyers and academic social scientists outscore CEOs and mayors in terms of cognitive ability. Medical doctors – a highly prestigious profession in Sweden associated with excellence – clearly show the highest cognitive scores of all. Academic economists and political scientists have the most years of education, rank second and third in cognitive scores, but have among the lowest leadership scores.

[Table 3 about here]

The patterns in the table make intuitive sense. Academics lack leadership but are smart, and as a result they accumulate the most years of education, but neither lead organizations nor make life-or-death decisions. Mayors and CEOs are marginally less smart, substantially less educated, but have higher leadership scores and, fittingly, do lead public and private organizations.

4.2 Elitism or Meritocracy?

The evidence presented so far shows strong positive selection of politicians, and increasingly so the higher the level of political attainment. This pattern could have two very different explanations. In one, Swedish politics is a meritocracy, selecting among the best and brightest. In another, Swedish politics is elitist, where heirs of rich families get privileged access to political power, and also more education and earning opportunities. Under the elitist explanation, the competence of politicians is a side effect, and does not play a preeminent role in selection.[9]

The elitist account would have two implications. On the one hand, if factors like wealth dominate ability, family background should matter greatly for political selection. On the other hand, conditional on family background, ability should matter little. We now address empirically these two implications.

[9] Against the elitist view, one could point to the facts that the education system in Sweden is entirely financed by the public sector, that admission to higher education is entirely based on high-school grades, and that education traditionally has been provided in roughly equal quality across the country.

Politicians and their siblings Figure 2 compares the distribution of competence traits of elected politicians with that of their siblings. To make this comparable to Figure 1 we also include the distribution in the full population. Clearly, elected politicians have markedly higher cognitive and leadership scores than their siblings, as well as higher residual ability.

We can also compare the *difference* in selection between politicians and siblings to the difference between politicians and the population. The difference vis-a-vis siblings in terms of leadership scores can account for 78 percent of the gap vis-a-vis the population. For cognitive scores the corresponding number is 70 percent, and for residual ability it is 74 percent. These numbers strongly indicate that ability, rather than family background, is the key selection criterion.[10]

[Figure 2 about here]

Figure 3 gives additional evidence that family background is less important than individual traits. It classifies politicians according to their percentile in the income distribution compared to the rest of the population within their own birth year and gender. By definition, the general population would display a perfect uniform distribution with a density of 0.01 for each percentile (as drawn by the gray line in the figure). The left graph shows that politicians are disproportionately drawn from higher income percentiles. But the distribution in the right graph for the politician siblings is very similar to the uniform distribution for the population. The fortunes of politicians thus appear related to their own ability traits rather than to family characteristics, as the latter would naturally extend to their siblings.

[Figure 3 about here]

Politicians and their parents We next examine the relevance of social background, and show that in the aggregate it does not matter much for political selection. We measure social background by parental income and occupational status, and show that politicians do not come disproportionately from elite backgrounds.

For parental income, we proceed as follows. For the politicians in the most recent election of our data, we find their parents' incomes and occupations in the earliest year(s) of our data. For about 80 percent of the politicians elected in 2010, we observe their father's income in 1979. In the analysis, we allow these fathers to be

[10] As an aside, the reader may wonder whether birth order is important. We find that politicians are more often first-born than their non-politician siblings. However, this fact does not explain the pattern in Figure 2, because all our ability measures are only marginally different for the first-born and later-born.

of any age, but the results remain the same when we instead use income observed when fathers were in the 35-45 age interval.

For each of the years 1979 and 2011, we use the full population data for individuals above 18 years of age to compute percentiles of the annual-earnings distribution within each gender and birthyear. We then compute the proportion of fathers (in 1979) and politicians (in 2011) with incomes within each percentile range. These proportions are shown in Figure 4. For elected politicians, the distribution is skewed to the right, showing a striking over-representation of high earnings relative to the population. But for fathers, the distribution has a uniform shape: politicians are almost perfectly representative of the population. (The corresponding figure for mothers' earnings distribution can be found in the Appendix, Figure A.1.)

[Figure 4 about here]

Politicians, CEOs, and medical doctors Again, it is valuable to compare the pattern for politicians with that of other elite professions. To do so, Figure 5 repeats the same exercise as Figure 4 for fathers of medical doctors (right graph) and CEOs of firms with 10-25 employees (left graph). These figures show that show that the 1979 earnings for fathers of doctors and CEOs are also much more skewed to the right than the earnings for fathers of politicians. Naturally, the earnings of the doctors and CEOs themselves (not shown) are also very much skewed to the right.

We can summarize the evidence from Figure 4 and Figure 5 in a different way. As measured by intergenerational earnings, social mobility into a political career seems to be high in absolute as well as relative terms, when compared to doctors and CEOs.

[Figure 5 about here]

Politicians in different parties The evidence so far concerns all elected politicians from all parties. In Figure 6, we replicate Figure 4 for the three largest parties in the municipal councils, the Social Democrats, the Conservatives and the Center (agricultural) party. As the three left graphs show, politicians in all parties come disproportionately from the top part of the income distribution, although more so in the Conservatives than in the Social Democrats or the Center party. In the three right graphs, however, we clearly see representation of different levels of earnings. High-income earners are over-represented among Conservative fathers, middle-income earners among Social-Democrat fathers, and low-income earners among Center-party fathers. The latter are likely to be small farmers (on average, 40 percent as opposed to 5 percent in other parties) with relatively low earnings.

[Figure 6 about here]

Figure 6 makes clear that different parties tend to represent different parts of the (parental) income distribution. The aggregation of these diverging party representations renders the almost perfect representation of parental incomes in Figure 4 for all politicians. Of course, this is not a coincidence but an illustration of the presumption that different parties will represent different interests – at least in polities with a multi-party system where the left-to-right dimension plays an important role.

Social class rather than income Although parental earnings are informative of social representation, they may only capture one part of social structure. Therefore, we supplement the data on parental income with data on parental social class whenever the relevant information is available. Figure 7 compares distributions of social class for politician parents and the population. The class division is based on occupation and corresponds closely to the EGP social-class scheme (Erikson and Goldthorpe 1992) which has been used by Lindgren, Oskarsson, and Dawes (2016) in their recent study of representation in Sweden. We define six classes as: (1) non-skilled manual workers, (2) skilled manual workers, (3) lower non-manual workers, (4) farmers, (5) intermediate non-manual workers, and (6) higher non-manual workers.[11] The data are again from 1979, and 54 percent of the politicians nominated in 2010 have a father for which we can define socioeconomic status.

[Figure 7 about here]

The pattern in Figure 7 corroborates our previous finding on earnings: politicians are highly representative of the population in terms of socioeconomic background. Farmers is the only social class that stands out as notably over-represented – something that reflects the historical role of the Center party. In addition, we see some under-representation of skilled manual workers.

An alternative interpretation of our findings on positive selection might be that meritocracy and elitism are not rival explanations, but one and the same. If individual competence is helped by parental investments in human capital, a strictly meritocratic system will favor elites. Meritocracy could then favor the competent within a family, but still be elitist across families. However, the finding that different social classes are evenly represented rejects this interpretation but favors our interpretation that Sweden's political system, on average, is both meritocratic and broadly representative.

[11] We are grateful to Martin Hallsten for sharing his STATA code with us. We are forced to drop the category of "self-employed" because of data constraints.

5 Selection Across Municipalities

Thus far, we have documented that politicians are positively selected from the Swedish population, and that this almost surely reflects meritocratic forces rather than elitism. Moreover, politicians represent social backgrounds proportionally. This does not imply, however, that politicians are positively selected in all 290 municipalities, or that all municipalities achieve even representation.

Selection indices defined To characterize municipal selection and representation, we compare the politicians in each municipal council to the general municipal population. We construct a simple selection index, as follows. For an ability or social-background variable x with K categories, we write the index as

$$S_x = \sum_{k=1}^{K} p_{k,c} k - \sum_{k=1}^{K} p_{k,m} k \ , \tag{2}$$

where $p_{k,c}$ is the proportion of council members in each category k, and $p_{k,m}$ is the corresponding proportion in the municipal population. We will use this index to gauge ability (when k is a competence score), as well as representation (when k is an ordered value tracking social background).

The resulting indices vary both across municipalities and within municipalities over time. For any trait where categories correspond to percentiles, the municipal population has a uniform distribution with an average percentile of 50. If politicians are positively (negatively) selected on competence in that municipality, they are drawn on average from percentiles higher (lower) than 50, so the selection index will be positive (negative). When it comes to the representation aspect of selection, a positive (negative) index instead reflects over-representation of higher (lower) parental incomes or social classes. Accordingly, an index of 0 corresponds to balanced representation.

Competence across municipalities Figure 8 shows the distribution of the political selection index for the three ability measures using municipality-level data for the 1991-2010 period. As indicated before, elected politicians are more able than the population according to all measures: elected politicians in the average municipality have, on average, cognitive scores ten percent higher than the population, with a similar pattern for the leadership score. Despite this positive average selection, we see considerable spatial variation with some municipalities even exhibiting adverse selection.

[Figure 8 about here]

Representation across municipalities Figure 9 plots the municipal distribution for our two representation measures. Some municipalities over-represent more privileged backgrounds, while others over-represent underprivileged ones. The distribution of the fathers-income selection index (to the right) is more or less centered at zero, as we would expect from Figure 4. The distribution of the fathers-social-class index (to the left) has mostly positive support, with modal values between 0 and 0.5. The key to this is the pattern in Figure 7, which shows over-representation of politicians with a farmer father (occupation value 4, through the Center party). This is more pronounced in smaller municipalities with smaller councils (as the Center party is strongest in those councils). As Figure 9 compares municipalities while Figure 7 compares individuals, the over-representation shows up in a stronger way in Figure 9.

[Figure 9 about here]

Is there a tradeoff? Figures 8 and 9 show a considerable dispersion across municipalities in ability and representation. It is natural to ask whether the two aspects of selection are correlated. In particular, we may suspect that a certain municipality that improves its selection in the ability dimension faces a cost of worse representation of lower socioeconomic backgrounds? If this were generally true, we should see a positive correlation between municipal competence and representation indices.

In Figure 10, we plot binned averages of our two representation indices (for father's income and social class) against the three ability indices (for cognitive score, leadership score, and residual ability). In each figure, we also provide the bivariate correlation (b) and the normalized relation in terms of standard deviations (*beta*). Positive values suggest positive relations between ability selection and overrepresentation of socioeconomically advantaged fathers.

[Figure 10 about here]

Overall, the plots suggest a very weak tradeoff between competence and social class (left column). The estimated slopes are all positive, but the slope coefficients are small. The strongest relationship, between father's social class and politicians' cognitive scores, suggests that a one standard deviation higher over-representation of higher social classes is associated with a 0.15 standard deviation higher cognitive score. For the two other competence measures, the relationships are much weaker.

Among the plots of competence and parental incomes (right column), we find the strongest relationship for residual ability. A 10 percentile increase in over-representation of parental income is associated with a higher average residual ability of 0.05 (with a corresponding *beta* estimate of 0.15). For the cognitive and leadership scores, the estimates are of even smaller magnitudes and not statistically significant.

Why is the tradeoff so flat? Why is the relationship between social representation and competence so weak? The flat tradeoff may be particularly surprising, given the measurement problems in disentangling innate ability from parental background. Enlistment concerns 18-year olds and, despite Sweden's comparatively uniform education system, parental background will likely shape measured ability via socialization and home resources. To shed further light on the lack of a meaningful trade-off, we investigate the extent to which selection on ability varies across parental socioeconomic groups.

[Figure 11 about here]

For each of our three ability measures, Figure 11 plots the difference in mean ability between elected politicians and the working population by the social class (occupation) of their fathers. The width of the bars corresponds to the share of politicians from each social class. As the bars indicate, positive selection on ability tends to be larger in lower social classes, which mitigates the ability cost of recruiting politicians from less favorable backgrounds. This is particularly the case for the cognitive score, where politicians from the lowest parental social classes are clearly more positively selected than those from the highest. These patterns help explain the flat tradeoff between ability and representation in Figure 10. Figure 12 shows similar patterns for politician vs. workforce ability by deciles of father's income.

One aspect of Figures 11 and 12 worth emphasizing is the remarkable stability of positive selection out of all father's social classes and income deciles.

What drives the variation across municipalities? One question is whether the selection of ability and representation, and the tradeoff between the two, are systematically related to observable municipal characteristics. We have looked for a number of economic, social and political drivers, including municipality size and location, level of income, income inequality, share of foreign-born, political majority, and level of political competition. We have not yet identified robust findings (results available upon request).

6 Selection Across Individuals

In this section, we begin by discussing the individual incentives to self-select into politics. To approach the data, we present a simple model, the comparative statics of which give us some guidance through a set of predictions. Informed by these, we look for some salient correlations. However, the equilibrium set of selected politicians depends not just on those who are willing to serve, but also on their party allowing a subset of the willing a slot on the party ballot. To shed light on this, we also discuss – theoretically as well as empirically – the screening by political parties.

6.1 A Simple Model of Political Selection

We write down and analyze a simple "Roy model" to shed light on self-selection into politics and the role of party screening.

Supply side – basic assumptions Consider a set of risk-neutral people, who have to decide whether to supply their services as leisure politicians. Each person is drawn from a continuous distribution jointly defined over two parameters: ability index Y (with typical element y) and index P (with typical element p) of intrinsic motives to serve. We assume these parameters are independent, with $P \in \left[0, \bar{P} > 0\right]$ uniformly distributed and $Y \in [0, \infty)$ distributed according to c.d.f. $G(y)$ with density $g(y)$.

Each person has a two-period horizon and there is no discounting. For simplicity, we assume that going into politics is a once-and-for-all choice in period 1.[12] Whoever does not go into politics earns y in the first period and expects to earn $\gamma y \geq y$ in the second. In other words, $\gamma \geq 1$ is a measure of the earnings-tenure profile (in the data, we allow γ to vary across occupations).[13]

Someone who offers to serve gets accepted to run and is elected to the municipal council with probability $q(y)$ – we consider different slopes of the screening $q(y)$ function when discussing the demand side. Elected politicians get non-pecuniary

[12] At the cost of introducing complexity, the model can be extended to include discounting as well as sequential decisions: a person who entered in period 1 can decide whether to stay or leave in period 2. Discounting adds notational complication only, while sequential decisions create more complex selection patterns which converge to those presented here as the parameter δ (to be defined shortly) goes to 1.

[13] This model of the supply side is related to those in Delfgaauw and Dur (2007), Francois (2000) and Dal Bó, Finan, and Rossi (2013), but among other differences it considers the distinct case of "leisure" politicians who do not give up immediate private sector earnings when entering public service, as well as the role of the experience premium γ.

benefits $\frac{p}{2}$ in each of the two periods. They must also give up some career: period-1 private earnings are y, but period-2 expected earnings are $(1-\delta)\gamma y$, where $\delta \in [0,1]$ shapes the future opportunity cost of current political involvement. The opportunity cost of politics is thus $\delta\gamma y$, which depends on general ability, private-career prospects, and how these are damaged by a political career

Some first-period council members are appointed mayors in the second period, in which case they earn political wage $w < \overline{Y}$ (and intrinsic benefit $\frac{p}{2}$). This happens with probability π.(in the data, we allow w and π to differ by municipality and party).

Cost-benefit calculation A person decides to become a politician when

$$(1+\gamma)y \leq (1-q(y))(1+\gamma)y + q(y)((1+(1-\pi)(1-\delta)\gamma)y + \pi w) + q(y)p.$$

After some algebra, this condition simplifies to

$$p + \pi(w - (1-\delta)\gamma y) \geq \delta\gamma y.$$

The intrinsic return to politics (the 1st term on the LHS) plus the probability of an income gain when becoming mayor (the 2nd term on the LHS) has to outweigh the opportunity cost of lost career prospects (the RHS).

The entry condition can be re-written as:

$$p \geq p(y) \equiv \pi((1-\delta)\gamma y - w) + \delta\gamma y. \tag{3}$$

Any type (y,p) on the "selection line" $p(y)$ is indifferent between entering politics and staying out. Those above (below) this line want to enter (stay out).

Comparative statics From the selection line defined in (3), we can derive the effect of a change in w as,

$$\frac{dp\,(y)}{dw} = -\pi < 0,$$

meaning that the selection line shifts down and the set of those willing to enter gets larger. As we prove in the Appendix, independent traits and a uniform distribution over P, imply that average ability must go up with any parametric change that shifts the selection line down in parallel fashion.

For parameter π, we get,

$$\frac{dp\,(y)}{d\pi} = (1-\delta)\,\gamma y - w,$$

which in general is ambiguous in sign. Here, the selection line pivots. As ability goes to zero, the derivative approaches $-w$, meaning that the line shifts down. But for very able types, it is positive if δ is low enough, and the selection line shifts up. Then, a higher π would reduce supply of high-ability types. But the data tell a different story, which can only be the case if δ is high enough that the line shifts down for all types (for any finite y, there exists δ close enough to 1 to render $\frac{dp(y)}{d\pi}$ unambiguously negative).

Finally, for parameter γ, (3) implies that,

$$\frac{dp\left(y\right)}{d\gamma} = \left(\pi + \delta\left(1 - \pi\right)\right) y > 0,$$

meaning a steeper earnings-tenure profile shifts the line up. This discourages entry, lowers average ability, and thus makes for worse supply.

We can summarize the comparative statics in the following (the Appendix gives a formal proof):

Proposition 1 *Suppose (p, y) are drawn from an independent distribution with P uniformly distributed on $\left[0, \bar{P} > 0\right]$ and Y distributed on $[0, \infty)$ according to $G\left(y\right)$. Then, maximum as well as average competence of people self-selecting into politics increase (weakly) with higher w and lower γ, and with higher π, if δ is high enough.*

Proof. See Appendix ■

This simple model of supply and its comparative statics resonate with the economic models of selection into politics discussed in the introduction, in that they point to clear material motives and opportunity costs as important drivers of self-selection. In addition, our model highlights intrinsic motives and dynamic career concerns.

Demand side Consider three types of screening of candidates that offer political service: (i) random selection (e.g., Athenian democracy, with election probability q unrelated to y); (ii) negative selection (e.g., cronyism, with $q'\left(y\right) < 0$); and (iii) positive selection (i.e., meritocracy, with $q'\left(y\right) > 0$).

Since Sweden is a party-based democracy, we assume that a party planner selects from the available candidate pool, anticipating voter demands. The main question is whether our earlier finding that elected politicians have higher ability than the average citizen means that parties screen in a positive way.

The answer is in the affirmative:

Proposition 2 *If the party has sufficiently good information on those who supply their services, the fact that elected politicians are more competent than the average citizen implies positive screening by parties. Then, the results in Proposition 1 extend to equilibrium outcomes.*

Proof. See Appendix. ■

To see this, note first that the term $q(y)$ does not affect the cost-benefit calculus of individuals. Since entry is invariant to party screening, we only need to keep track of the entry condition to characterize the candidate pool. Suppose the planner observes candidate types (y, p) perfectly. We can then show that both random and negative screening must lead to politicians less competent than the average citizen, leaving positive screening as the only remaining alternative.[14]

Consider random selection. Given selection line $p = \pi((1-\delta)\gamma y - w) + \delta\gamma y$, the relatively competent self-select out of politics, and expected candidate ability (denote it $E(y_A)$) must be worse than the average ability $E(y)$ in the population. A bit more formally, (and abstracting from the fact that $\gamma \geq 1$) note that the entry condition implies $p(y) \to -\pi w$ if $\gamma \to 0$, such that all citizens enter and $E(y_A) \to E(y)$. But as shown above, $E(y_A)$ decreases in γ. This means that when we raise γ away from zero, $E(y_A)$ must dip below $E(y)$.

As random screening implies that the average quality of selected politicians is worse than that of the population, the result for negative screening follows immediately.

Because of positive screening, our comparative statics of supply characterize those who are selected into parties and elected. This is easy to see if the party observes types (p, y) perfectly.[15] Any party that values competence will select individuals with the highest available ability, i.e., those with type $\left(\bar{P}, \frac{\bar{P}+w\pi}{(\pi(1-\delta)+\delta)\gamma}\right)$. Thus, any change in (γ, π, w) that shifts the line down will (weakly) increase the average and top quality not just among those willing to enter, but also among those elected. Thus, the competence of politicians is weakly increasing in w and π (when δ is high enough), and weakly decreasing in γ.

[14] While our conclusion of positive screening relies on the assumptions of our model, specifically the one on the distribution of traits, it is corroborated by evidence presented below on the screening role of parties.

[15] The argument can be extended to the case when the party observes types imperfectly at the cost of some additional notation and algebra.

6.2 Evidence on Self-selection

The model in the previous subsection helps us identify some drivers of self-selection. Its comparative statics suggest that mayoral wages and appointment probabilities as well as earnings-tenure profiles should be systematically related to the competence of politicians. In this subsection, we check whether the correlations in the data are consistent with these predictions.

Mayor earnings Our model predicts higher monetary remuneration to attract more competent politicians. To explore this prediction, we focus on the salary of mayors, the only (or one of few) full-time paid position(s).[16] Mayor salaries vary substantially across municipalities. In 2011, their average annual earnings was 632,400 SEK (about 79,044 USD), with a standard deviation of 213,000 SEK.

To relate the value of a this wage to income opportunities in the municipality, we normalize the mayor's annual earnings by average earnings among all inhabitants above 18 years of age. This approximates the model's w, the material payoff to the position as mayor.

We consider a sample of all local parties that ever appointed a mayor in the period 1991-2010. Because the probability of becoming mayor varies by rank on the municipal party list – first-ranked politicians being the most likely, second-ranked being next in line, and so on – we select the top-five people from every electoral ballot and create five samples, one for each list rank.

Then, we use OLS to estimate

$$Q_{i,r,m,t} = \alpha_t + \beta_r w_{m,t} + \varepsilon_{i,r,t}, \tag{4}$$

where $Q_{i,r,m,t}$ is one of our three ability measures for politician i with list rank r, in municipality m in election period t, α_t is an election-period fixed effect, and $w_{m,t}$ is the (normalized) mayoral wage in municipality m and election period t. The coefficient of interest is β_r, the linear relation between the mayor's relative wage and the selection of politicians for rank r. If a higher salary attracts higher-ability individuals, β_r should be positive. Moreover, if high earnings attract high-ability candidates to seek positions with a higher probability to become mayor, β_r should be higher for $r = 1$.

Figure 13 plots our estimate of interest from equation (4), from our five r-samples. The estimates for the cognitive and leadership scores (left graph) show a positive and

[16] We could also add the prospects of a promotion to a position i the national parliament. Because the probability of this event is so low (on the order of 0.5%), this would not change any of the results to be shown here.

significant relationship only for the selection into top rank. For the remaining ranks, the estimates are smaller in size, close to zero, and not statistically significant. The estimates for residual ability (right graph) show positive and significant estimates for $r = 1, 2, 4$, with the largest point estimate for the top rank.

[Figure 13 about here]

The correlations in Figure 13 are clearly consistent with our model, where brighter career prospects in the form of a higher mayor's wage draw more able people into politics.

Appointment probabilities In the model, π is the probability that an elected politician is promoted to mayor. This probability varies with the political status of the party. Some parties are in a strong majority position, making their appointing the mayor highly probable. Other parties are small in size and/or belong to the opposition bloc, making their appointing the mayor highly unlikely. We thus classify parties into categories depending on the political-career opportunity they afford.

To measure the appointment probability of a party, we simply calculate the proportion of election periods between 1982 and 2010 that it appointed the mayor. Specifically, we divide parties into four groups: (1) 100 % probability, (2) 51-99% probability, (3) l-50% probability, and (4) 0 % probability.

We compare competence selection indices across these four categories. We want to know: (i) does type (1), with the highest probability of appointing the mayor, stands out in terms of positive selection? (ii) do parties of other types, especially type (4), still show positive selection of politicians?

Table 4 shows average competence selection indices within each party category, for all elected politicians and for the person who tops the party's electoral ballot. On question (i), we find that parties that are sure to appoint the mayor indeed have a better selection of their top-ranked politicians.[17] Hence, the material career prospects do seem to matter for positive selection, as our model suggests.

[Table 4 about here]

Indirect evidence on intrinsic motives On question (ii), we find no evidence of adverse or neutral selection for the rank and file in parties with a small (or zero) probability of promotion. The average representative is as qualified in the party

[17] Majority parties have larger party delegations on average, which means that the average competence among the rank and file is pulled down by moving further down the competence distribution.

category with the weakest career prospects as in the party categories with better prospects.

In terms of our model, this suggests that material motives tell only part of the self-selection story and that intrinsic motives must also play an important role. To see this, recall the entry condition $p(y) \geq \pi((1-\delta)\gamma y - w) + \delta\gamma y$. When the probability of appointment $\pi \to 0$, this can obviously not be fulfilled for any level of y unless $p > 0$.

Earnings-tenure profiles The third prediction from our simple model is that selection is less positive for people whose private career has higher returns to experience – i.e., a higher value of parameter γ.

To shed light on this, we compute earnings-tenure profiles for different occupations in two ways. One builds on a categorization of easily identifiable education types, which cover roughly 70% of the working-age population. The other way builds on sectors of employment, the same sectors that go into the estimation of the Mincer equation in (1) underlying our measure of residual ability. As in that estimation, we divide the people in each sector into two groups, one with tertiary education and one without. This categorization covers the whole working-age population but does not lend itself to easy labeling as the first method.

For each labor-market segment, we first compute a proxy for γ, as the average rate of (nominal) earnings growth over the course of the sample. Then, we compute separate selection indices, like the one in (2), for our three ability measures and politicians relative to non-politicians in each different labor-market segment. Finally, we plot these selection indices against the estimated earnings-tenure profiles for each occupation.

The results are displayed in Figure 14, where each row of plots shows a specific ability selection index – from top to bottom, the leadership score, the cognitive score, and residual ability. The columns apply to a specific labor-market division: educational categories to the left and employment sectors to the right.

[Figure 14 about here]

As is evident from the figure, the data are consistent with the prediction from our simple model. Five graphs out of six show a downward-sloping relation, meaning that politicians in occupations with higher earnings growth (higher γ) are less positively (in some cases even negatively) selected than those in occupations with lower earnings growth.

6.3 Evidence on Party Screening

The model in Subsection 6.1 suggests that our findings in Section 4 – that the elected are more positively selected than the nominated, and mayors even more positively selected – indicate that parties engage in positive candidate screening. This subsection provides some evidence for such screening.

Background Parties screen candidates in different ways. One is to provide an arena where politicians can compete in coming up with good arguments and policy proposals. Such competition may well result in positive selection if more able politicians win out in the tournament and climb to the top of the party. Alternatively, party constituencies (e.g., the youth branch, the female branch, associated unions, etc.) can select and promote the more able to higher positions in the party list.

As mentioned in Section 2, qualitative work in political science suggest that political parties in Sweden actively screen and promote candidates. While we do not strive to identify the exact mechanisms, we would like to present at least some evidence that parties gradually promote the more competent to higher positions. This would help us further understand the pattern documented in Section 4.

Selection and list rank To do so, we consider all party lists in all municipalities, electoral periods and parties. From these, we compute an ability index for all candidates who hold a certain list rank, for each list rank between 1 and 8. The graphs in the left column of Figure 15 show how ability varies by list rank for the cognitive score, the leadership score, and residual ability. All measures show a more or less steady decline by list rank, with the clearest pattern for residual ability. In particular, the top-ranked politician has significantly higher ability than every other rank for all three competence measures.

[Figure 15 about here]

Parties thus seem to screen and promote more able people towards progressively higher positions on their ballots. Because of the interplay between party screening and self-selection, improvements among those that self-select into political service will translate into higher equilibrium competence of elected politicians. Healthy political parties – that can offer positive screening to society – thus appear as an important component of democracy.

Selection and political competition If parties screen and also want to win elections, then stiffer political competition should give parties stronger incentives to screen their candidates for ability. To see if this is the case, we partition municipal elections by the degree of political competition between the left and center-right political blocks. Specifically, we measure competition in each municipality and election by the difference between the vote shares of the two blocs in the last three elections.

The right column of graphs in Figure 15 shows that politicians in municipalities with political competition above the median are indeed more positively selected than those below the median at all party list ranks. This is further (indirect) evidence for the screening role of parties.

7 Conclusion

Research in political economics offers theoretical arguments and empirical evidence for the notion that leaders matter, and that societies benefit from an able and broadly representative leadership. While democracy may better suited than other political systems to promote representation, it is not clear that it can deliver a representative *and* able leadership. In this paper, we analyze political selection with regard to ability and representation in Sweden, a paradigmatic advanced democracy. We use a unique data set, with rich information on ability traits and social background for the universe of municipal politicians and the entire Swedish population. Based on these data, we uncover four facts: (1) Politicians are strongly positively selected for all competence measures; (2) Representation of social background, whether measured by intergenerational earnings or social class, is very even; (3) There is at most a weak tradeoff in selection between ability and representation; (4) Individual motives matter in selection, as does screening by political parties.

Although we cannot extrapolate these findings to the rest of the world, the facts we uncover are important in that they alleviate a concern that political systems encouraging broad representation select mediocre leaders. Still, some of the patterns in the data may be quite specific to Swedish political (and societal) institutions. Data permitting, it would thus be very interesting to carry out a comparative analysis of other countries with similar or dissimilar political systems.

Our findings suggest a need to rethink models of political selection and recruitment. Standard models of candidate entry, which focus mainly on the material motives of holding office, cannot explain positive selection. The data seem to support a richer view, where intrinsic as well as material motives shape entry decisions. Also, political parties clearly play an active role in selecting candidates, and our results suggest that candidate ability is an important element in this calculus. But

the precise criteria and processes used internally in party organizations is a topic in dire need of further empirical and theoretical work.

Future work should also extend the analysis in several dimensions. There is room for more comprehensive modelling to interpret and better guide the empirical work especially on the roles parties and voters play in selection. In addition, future research should attempt to quantify intrinsic motivation, and its impacts on selection and performance. Finally, it would be valuable to study how the competence and representativeness traits of political leaders leave a mark on policy outcomes.

References

Beath, A., F. Christia, G. Egorov, and R. Enikolopov (2014). Electoral Rules and the Quality of Politicians: Theory and Evidence from a Field Experiment in Afghanistan. Working Paper 20082, National Bureau of Economic Research.

Besley, T. and S. Coate (1997). An Economic Model of Representative Democracy. *The Quarterly Journal of Economics 112*(1), 85–114.

Besley, T., O. Folke, T. Persson, and J. Rickne (2013). Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden. Working Paper Series 985, Research Institute of Industrial Economics.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011). Do Educated Leaders Matter? *Economic Journal 121*(554), F205–227.

Besley, T., R. Pande, and V. Rao (2005). Political Selection and the Quality of Government: Evidence from South India. Center Discussion Papers 921, Economic Growth Center, Yale University.

Besley, T. and M. Reynal-Querol (2011). Do Democracies Select More Educated Leaders? *American Political Science Review 105*(3), 552–556.

Carlstedt, B. (2000). *Cognitive Abilities. Aspects of Structure, Process and Measurement*. Dissertation, University of Gothenburg.

Caselli, F. and M. Morelli (2004). Bad Politicians. *Journal of Public Economics 88*(3-4), 759–782.

Chattopadhyay, R. and E. Duflo (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica 72*(5), 1409–1443.

Dahl, S. (2011). *Efter folkrörelsepartiet: Om aktivism och politiska kursomläggningar i tre svenska riksdagspartier (Life after the popular parties: About activism and political re-directions of three parties with parliamentary representation)*. Dissertation, Stockholm University, Stockholm Studies in Politics 140.

Dal Bó, E., P. Dal Bó, and J. Snyder (2009). Political Dynasties. *Review of Economic Studies 76*(1), 115–142.

Dal Bó, E., F. Finan, and M. Rossi (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics 128*(3), 1169–1218.

Delfgaauw, J. and R. Dur (2007). Signaling And Screening of Workers' Motivation. *Journal of Economic Behavior & Organization 62*(4), 605–624.

Djerf-Pierre, M. and B. Niklasson (2010). Vem vill ha jämställdhet? (Who wants gender equality?). In S. Holmberg and L. Weibull (Eds.), *Det våras för politiken: trettiotvå artiklar om politik, medier och samhälle (Springtime for politics: thirty-two articles about politics, the media, and society.* Oxford University Press.

Erikson, R. and J. Goldthorpe (1992). *The Constant Flux: a Study of Class Mobility in Industrial Societies*. Clarendon Press.

Erlingsson, Gissur , P. M. and R. hrvall (2012). Den motvilligt engagerade altruisten (The Reluctantly Engaged Altruist). *Statsvetenskaplig Tidskrift (Swedish Political Science Review) 114*(2), 185–205.

Ferraz, C. and F. Finan (2009). Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance. NBER Working Papers 14906, National Bureau of Economic Research, Inc.

Folke, O., T. Persson, and J. Rickne (2015). Preferential Voting, Accountability and Promotions into Political Power: Evidence from Sweden. *American Political Science Review*, Forthcoming.

Francois, P. (2000). 'Public Service Motivation' as an Argument for Government Provision. *Journal of Public Economics 78*(3), 275–299.

Fredriksson, J. (2001). Vgen till politiken. (the road to politics). In *Att vara med på riktigt demokratiutveckling i kommuner och landsting (To participate for real - democratic developments in municipalities and counties).* SOU 2001:48, Swedish Government White Paper.

Freidenvall, L. (2006). Vägen till Varannan damernas. Kvinnorepresentation, kandidaturval och kvotering i svensk politik 1970-2002 (The road to "zippered lists". Female representation, candidate choice and quotas in Swedish politics 1970-2002). Technical report, Stockholms Universitet.

Gagliarducci, S. and M. D. Paserman (2012). Gender Interactions within Hierarchies: Evidence from the Political Arena. *Review of Economic Studies 79*(3), 1021–1052.

Galasso, V. and T. Nannicini (2011). Competing On Good Politicians. *American Political Science Review 105*(1), 79–99.

Jia, R., M. Kudamatsu, and D. Seim (2015). Political Selection in China: The Complementary Roles of Connections and Performance. *Journal of the European Economic Association 13*(4), 631–668.

Jones, B. F. and B. A. Olken (2005). Do Leaders Matter? National Leadership and Growth Since World War II. *The Quarterly Journal of Economics 120*(3), 835–864.

Karlsson, D. (2001). Ny som förtroendevald i kommuner och landsting (new to local politics). In *Att vara med på riktigt demokratiutveckling i kommuner och landsting (To participate for real - democratic developments in municipalities and counties)*. SOU 2001:48, Swedish Government White Paper.

Lindgren, K.-O., S. Oskarsson, and C. Dawes (2016). Can Political Inequalities be Educated Away? Evidence from a Swedish School Reform. *American Journal of Political Science*, Forthcoming.

Lundqvist, H. (2013). Is it Worth it? On the Returns to Holding Political Office. Working Papers 2013/14, Institut d'Economia de Barcelona (IEB).

Merlo, A., V. Galasso, M. Landi, and A. Mattozzi (2010). The Labor Market of Italian politicians. In T. Boeri, A. Merlo, and A. Prat (Eds.), *The Ruling Class: Management and Politics in Modern Italy*. Oxford: Oxford University Press.

Messner, M. and M. K. Polborn (2004). Paying Politicians. *Journal of Public Economics 88*(12), 2423–2445.

Meyersson, E. (2014). Islamic Rule and the Empowerment of the Poor and Pious. *Econometrica 82*(1), 229–269.

Myerson, R. (1993). Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems. *American Political Science Review 87*(4), 856–869.

Nielsen, P. (2001). P och av om uppdragsvillighet, rekrytering och avhopp i den kommunala demokratin (on and off motivation for public office, recruitment and quits in the municipal democracy). In *Att vara med på riktigt demokratiutveckling i kommuner och landsting (To participate for real - democratic developments in municipalities and counties)*. SOU 2001:48, Swedish Government White Paper.

Niklasson, B. (2005). Contact Capital in Political Careers. Gender and Recruitment of Parliamentarians and Political Appointees. Technical report, University of Gothenburg.

Nilsson, P. J. (2014). Alcohol Availability, Prenatal Conditions, and Long-term Economic Outcomes. *Journal of Political Economy, forthcoming*.

Öhberg, P. (2010). Politiker med karrirambitioner en omjlig sjlvklarhet (Politicians with Career Ambitions A Contradiction. Doctoral dissertation, Gothenburg University.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press.

Osborne, M. and A. Slivinski (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics 111*(1), 65–96.

Pande, R. (2003). Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India. *American Economic Review 93*(4), 1132–1151.

Persson, T. and G. Tabellini (2003). *The Economic Effects of Constitutions*. MIT Press.

Powell, G. B. (2000). *Elections as Instruments of Democracy - Majoritarian and Proportional Visions*. New Haven, Conn.: Yale University Press.

Soininen, M. and N. Etzler (2006). *Partierna nominerar: Exkluderingens mekanismer etnicitet och representation (Political parties nominate: Mechanisms of exclusion ethnicity and representation)*. Stockholm, Sweden.: Swedish Government Official Report, SOU, 2006:53s.

SOU:68 (2007). Ett decennium med personval (Lessons from a Decade of Preference Voting). Swedish Government White Paper.

Ståhlberg-Carlstedt, B. and P. Sköld (1981). Inskrivningsprov 80 - Testmanual (Enlistment Test 80 - Testing Manual). Report CH 55001-H7, National Defense Research Establishment Report.

Statistics Sweden (2010). Åttapartivalet 2010, Allmänna valen Valundersökningen (The eight parties election 2010, general election, electoral study). Official Statistics Sweden Report.

Taagepera, R. and M. Shugart (2000). *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, Conn.: Yale University Press.

Teorell, J. (2000). Resource Model of Social Capital: Networks, Recruitment and Political Participation in Sweden. conference presentation, European Consortium for Political Research.

Tillmann, P. (2014). Entry into Electoral Races and the Quality of Representation. Unpublished, University of Rochester.

Washington, E. L. (2008). Female Socialization: How Daughters Affect Their Legislator Fathers. *American Economic Review 98*(1), 311–32.

8 Figures

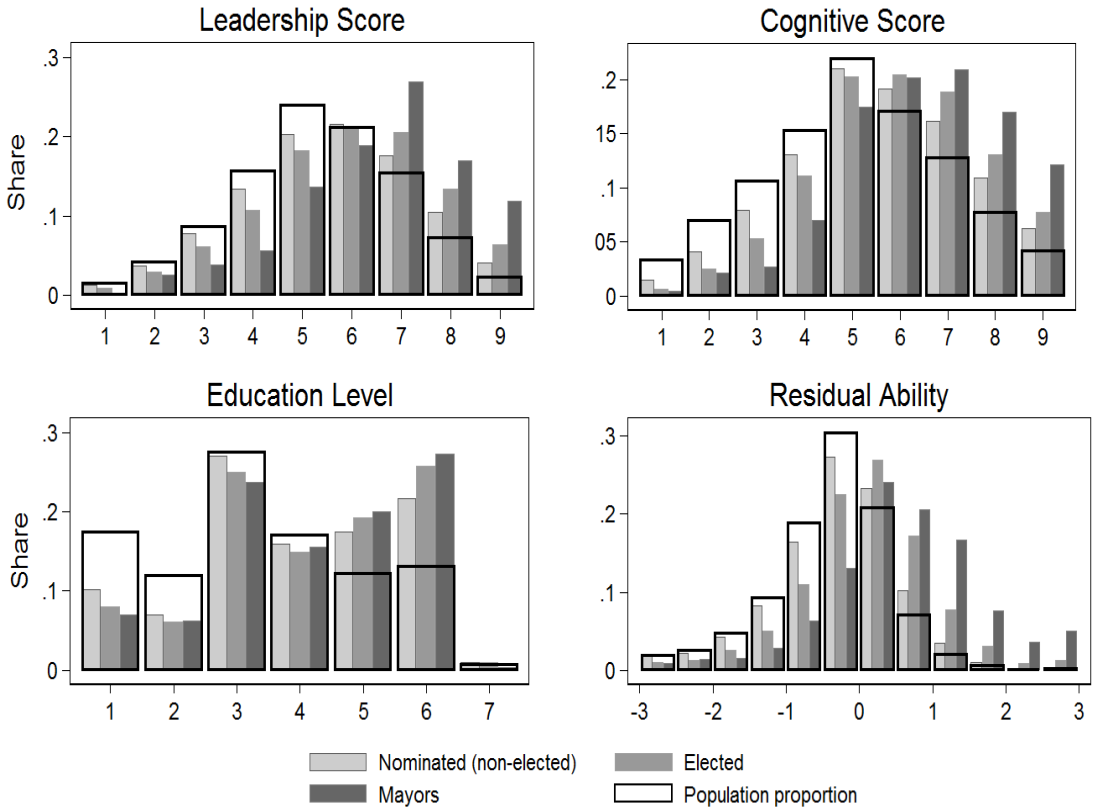

Figure 1: Distribution of ability measures in the population and among politicians

Notes: The figure shows comparisons of the distribution of the ability variables among the Swedish working-age population (18-72) and three categories of politicians: nominated (non-elected), elected, and mayors. The two enlistment scores are shown in the top figures; the bottom left figure refers to the level of education, and the bottom right the residual ability measure. Level of education is coded into seven groups based on the formal categorization of Statistics Sweden, namely less than 9 years, 9 years of primary education, 2-year secondary education, 3-year secondary education, tertiary education (less than three years), tertiary education (at least three years) and research degree (licenciate or Ph.D.) All figures are created with pooled individual level data for election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006 and 2010). For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts. Based on a Kolmogorov-Smirnov test, we can reject that the distributions are the same for each panel and for every pair of categories.

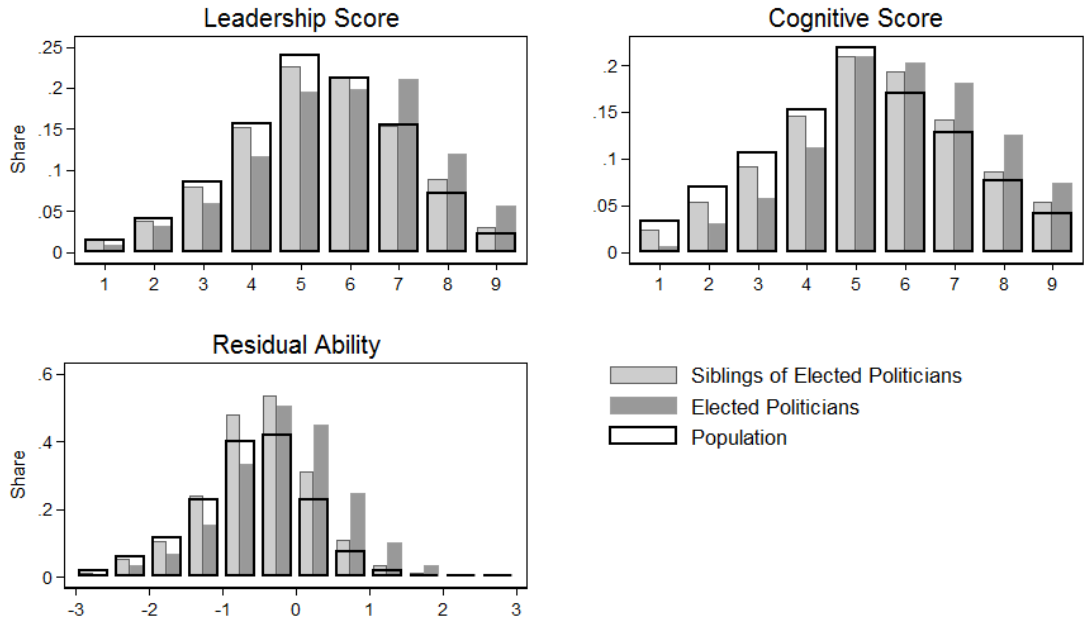

Figure 2: Distribution of ability measures among elected politicians, their siblings, and the population

Notes: The figure compares the distributions of our ability variables for elected politicians and their siblings. The two upper graphs show the distributions for the leadership score (left) the cognitive score (right) and the lower graph shows residual ability. The proportion of elected politicians with each score are represented by the dark gray bars, while the light gray bars show the proportions among the siblings. Only elected politicians with at least one sibling are included. The wider, hollow, bar graphs indicate the population proportions for each score. All figures are created with pooled individual level data for election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006 and 2010). For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

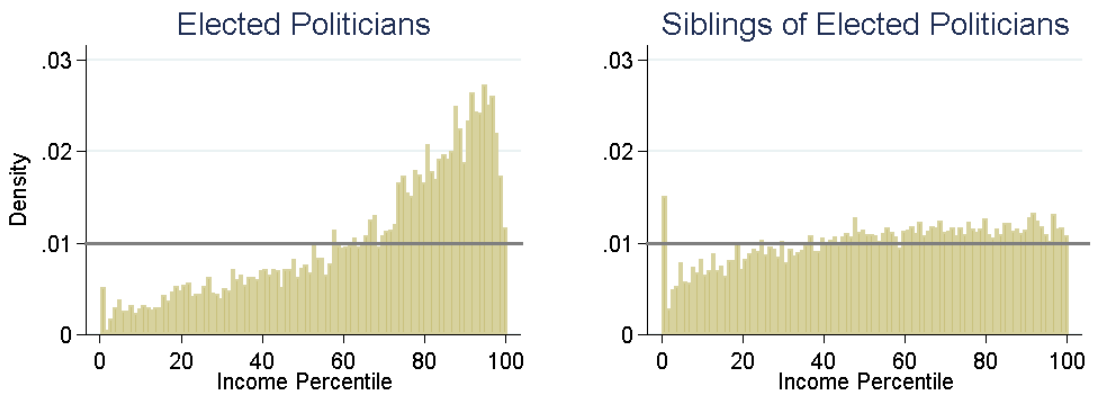

Figure 3: Distribution of elected politicians and their siblings across the percentiles of population income, 2011

Notes: The figure compares the distribution of annual labor incomes of elected politicians (left) to that of their siblings (right). Data from year 2011 for the working-age population (18-72) was used to compute the percentiles of annual earnings in the population. The proportion of politicians and siblings in each percentile bracket is shown in the histograms. Only elected politicians with at least one sibling are included.

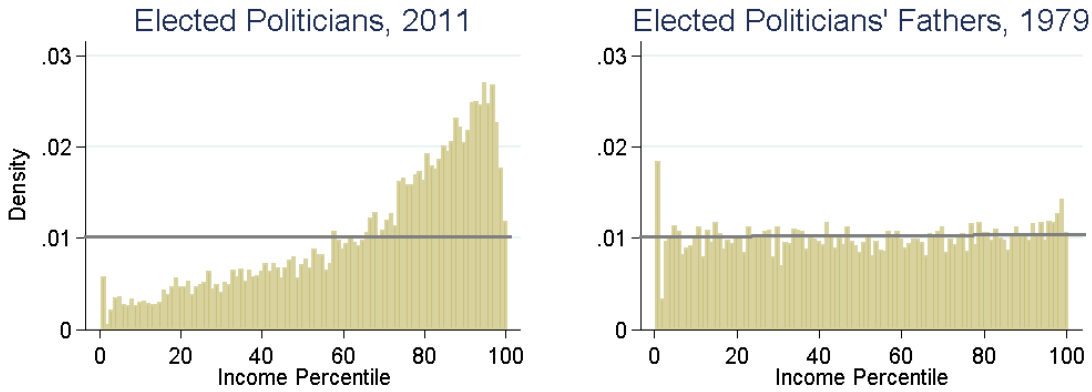

Figure 4: Distribution of elected politicians and their fathers across the percentiles of population income

Notes: The figure shows the distribution of elected politicians (left) and their fathers (right) across the percentiles of the Swedish income distribution. The income percentiles are calculated by birth year and gender. Data from year 1979 and 2011 were used to compute the percentiles of annual earnings in the working-age population (18-72) in these two years, respectively. The proportion of individuals who fall into each percentile bracket is shown in the histograms. Fathers are only included if they are of working age in year 1979 (18-72), and politicians are excluded if we cannot find an earnings observation for their father in that year.

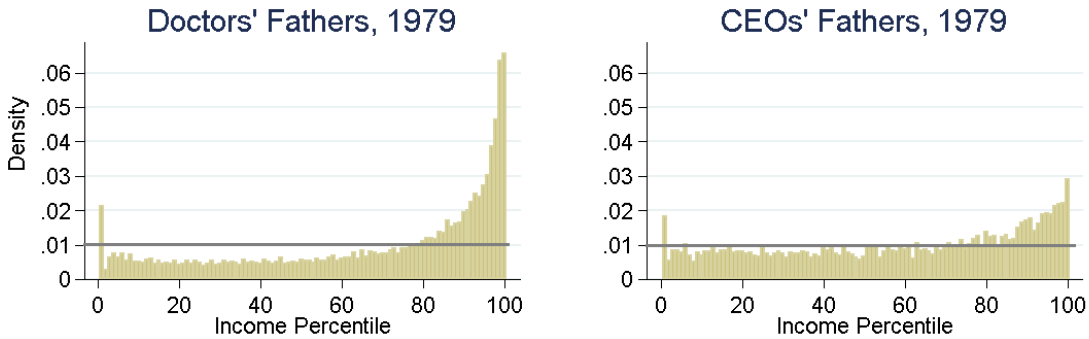

Figure 5: Distribution of the fathers of both medical doctors and CEOs across the percentiles of population income, 1979

Notes: The figure shows the distributions of the annual labor earnings in 1979 of the fathers of medical doctors (left) and CEOs (right). The income percentiles are calculated by birth year and gender. Data from year 1979 and 2011 was used to compute the percentiles of annual earnings in the working-age population (18-72) in these two years, respectively. The proportion of individuals who fall into each percentile bracket is shown in the histograms. Fathers are included if they are working age (18-72), and CEOs and medical doctors are excluded if we cannot find an earnings observation for their father in that year.

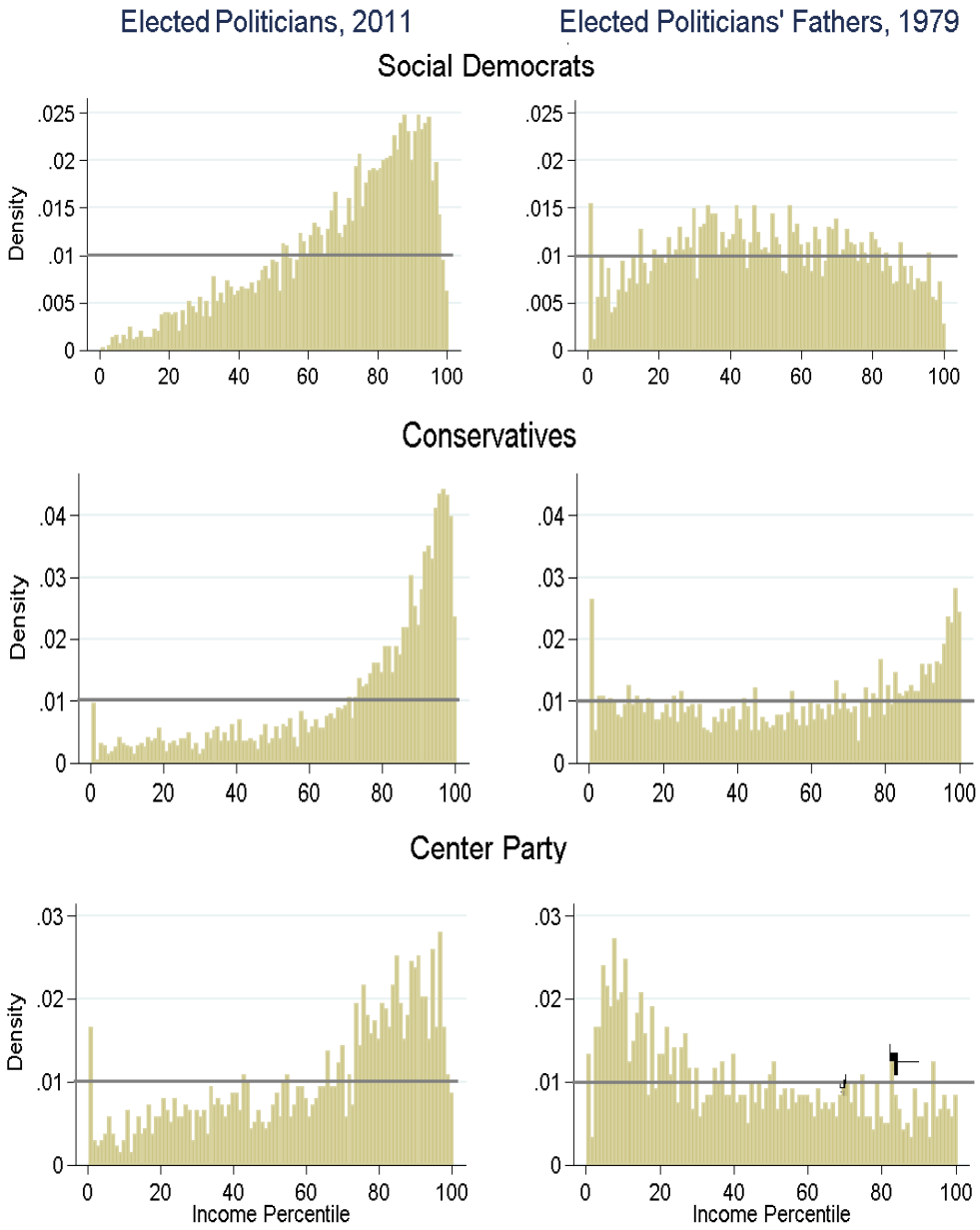

Figure 6: Distribution of elected politicians and their fathers across the percentiles of population income

Notes: The figure shows distributions of elected politicians (left) and their fathers (right) across the percentiles of the Swedish income distribution. The income percentiles are calculated by birth year and gender. The top (middle/bottom) figure includes politicians elected to a municipal assembly seat for the Social Democrats (Conservative/Center) party. See the notes for Figure 4 for details on the data used.

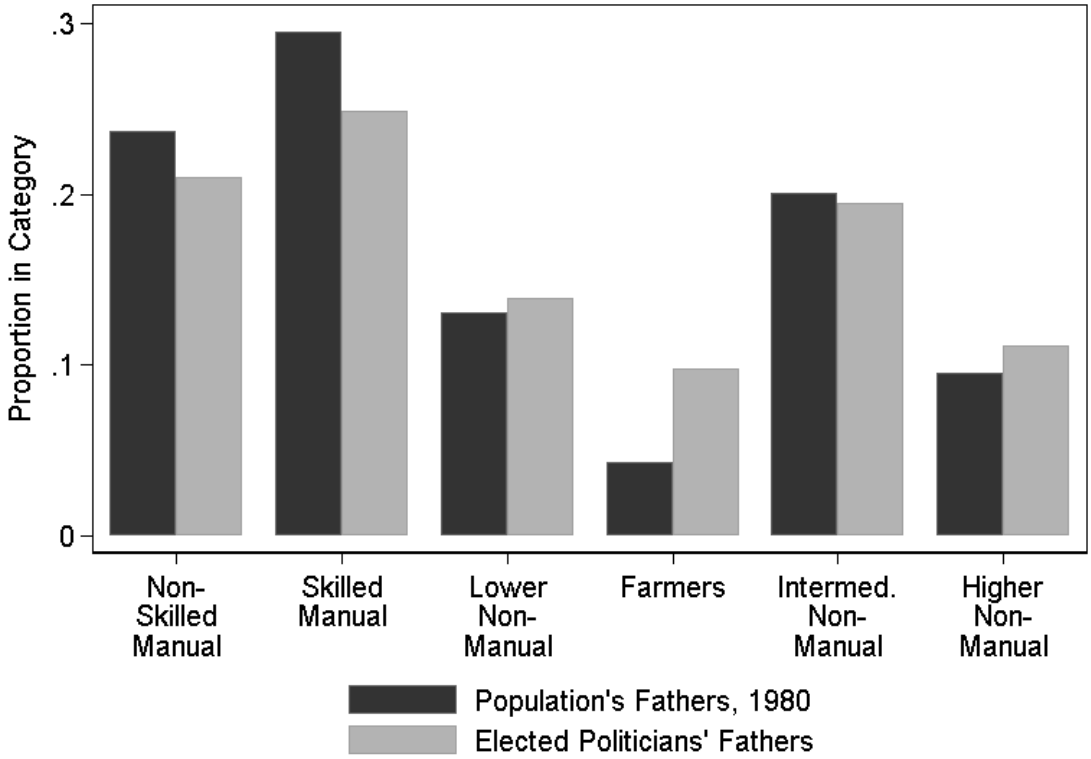

Figure 7: Distribution of parental social class of elected politicians and the population

Notes: The figure shows the distributions across six social classes of elected politicians and the population. The height of the bars shows the proportion of individuals in each social class. Elected politicians are measured in year 2010. Parents' social class is measured in 1979. Only politicians for whom we can identify a father or a mother with a non-missing social class are included. In cases where we can identify the social class of both parents, only the one with the highest social class is included in the figure.

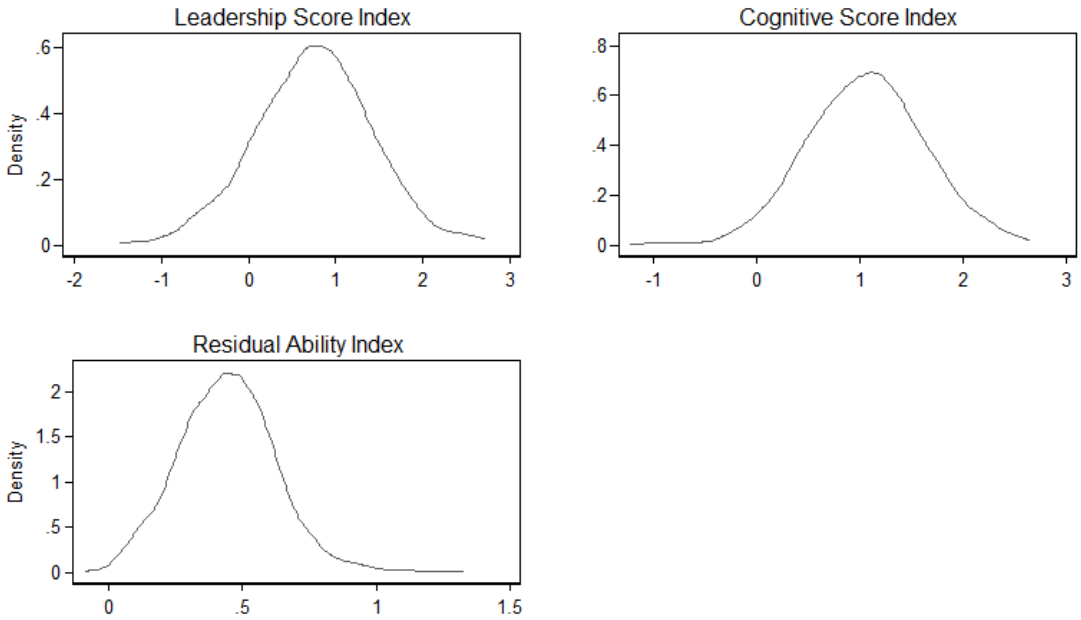

Figure 8: Distribution of ability selection indices for elected politicians

Notes: The figure shows the distribution of the political selection index for our three ability variables across municipalities and time. The unit of observation is the municipality and election period. The distributions for the leadership score and cognitive score are shown in the upper graphs, and the distribution for residual ability in the lower graph. The selection index is computed by subtracting from the average competence measure among elected politicians from the average of the same competence measure in the working-age population in the same municipality and election year. The data includes all election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006 and 2010).

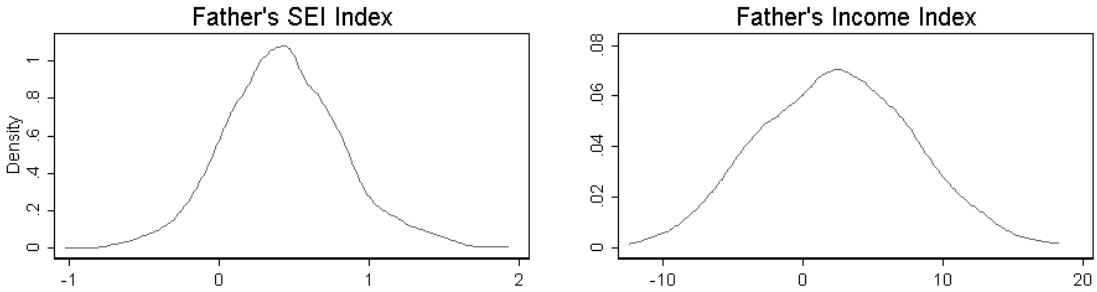

Figure 9: Distribution of representation indices across municipalities

Notes: The figure shows the distribution of two representation indices for politicians' parents across municipalities and time. The unit of observation is the municipality and election period. For politicians elected to a municipal council in the 2000s (2002, 2006, or 2010) we measure their parents' social class in 1980, and their parents annual earnings percentile by gender and birth year in 1979. Social classes are given the values 1-6 as: (1) Non-skilled manual; (2) Skilled manual; (3) Lower non-manual; (4) Farmer; (5) Intermediary non-manual; and (6) Higher non-manual. For each individual politician, we only include the parent in the highest annual earnings percentile, or social class (either the mother or the father). We then calculate the representation indices for social class (left) and income (right), as the average among politicians' parents minus the average of the parents of the working-age population (18-72) in the same municipality.

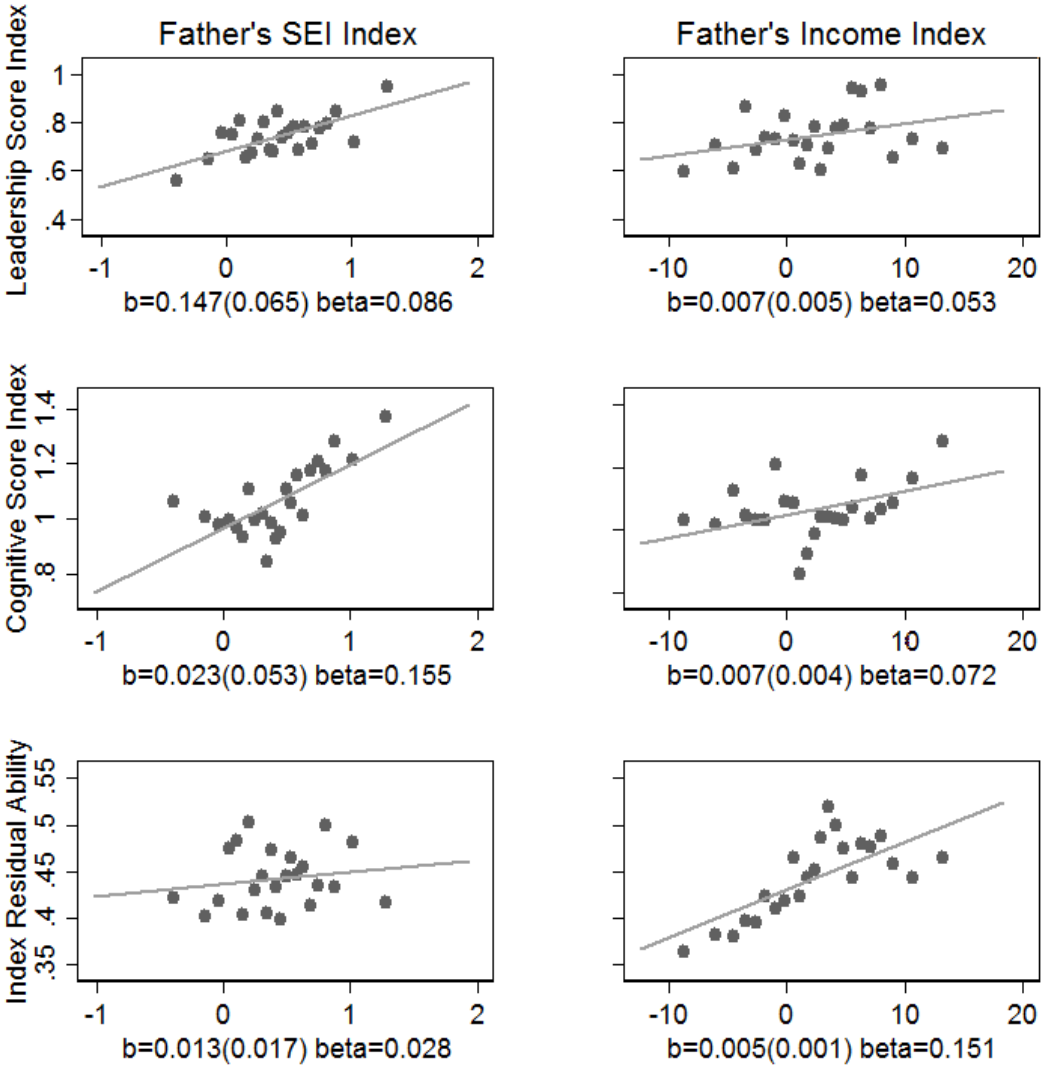

Figure 10: Correlations between municipal indices of representation and selection

Notes: The figure shows the relationships between competence selection indices (y-axis) and representation indices (x-axis). The unit of observation is the municipality and election period. Each dot in the scatter plots corresponds to the average among 1156 municipality-election values. The regression line shows the estimated slope coefficient from an OLS regression of the selection index on the representation index. The slope coefficient is displayed below each graph. Standard errors are reported in parenthesis, and *beta* is the normalized relationship in terms of standard errors. The data include all elections from 1998-2010. For the leadership and cognitive scores, data is restricted to men in the 1951-1980 cohorts.

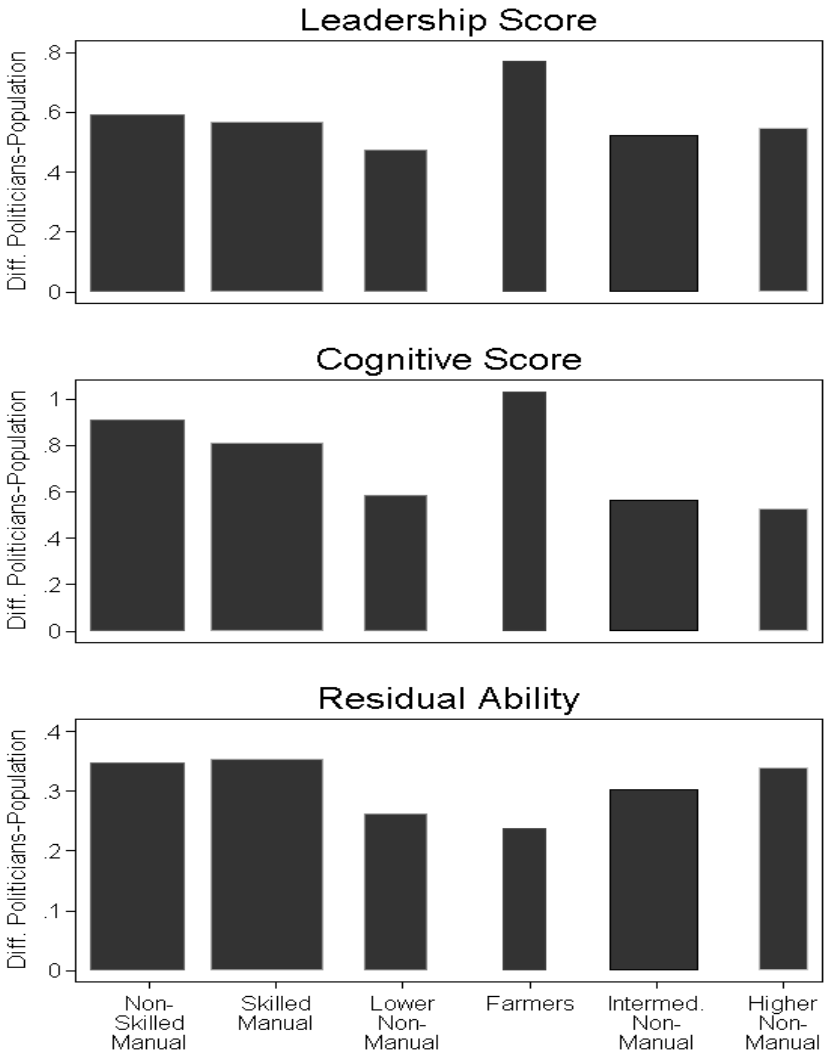

Figure 11: Difference in average ability measures between politicians and general population by father's occupational class

Notes: The figure shows differences in mean ability between politicians elected to a municipal council seat and the working-age population by the social class of each individual's father (x-axis), measured in the year 1980. The width of each bar represents the share of politicians whose father belongs to that particular social class. The data are pooled for all politicians elected in the 2000s (2002, 2006 and 2010), and the general population is also sampled in these same years. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

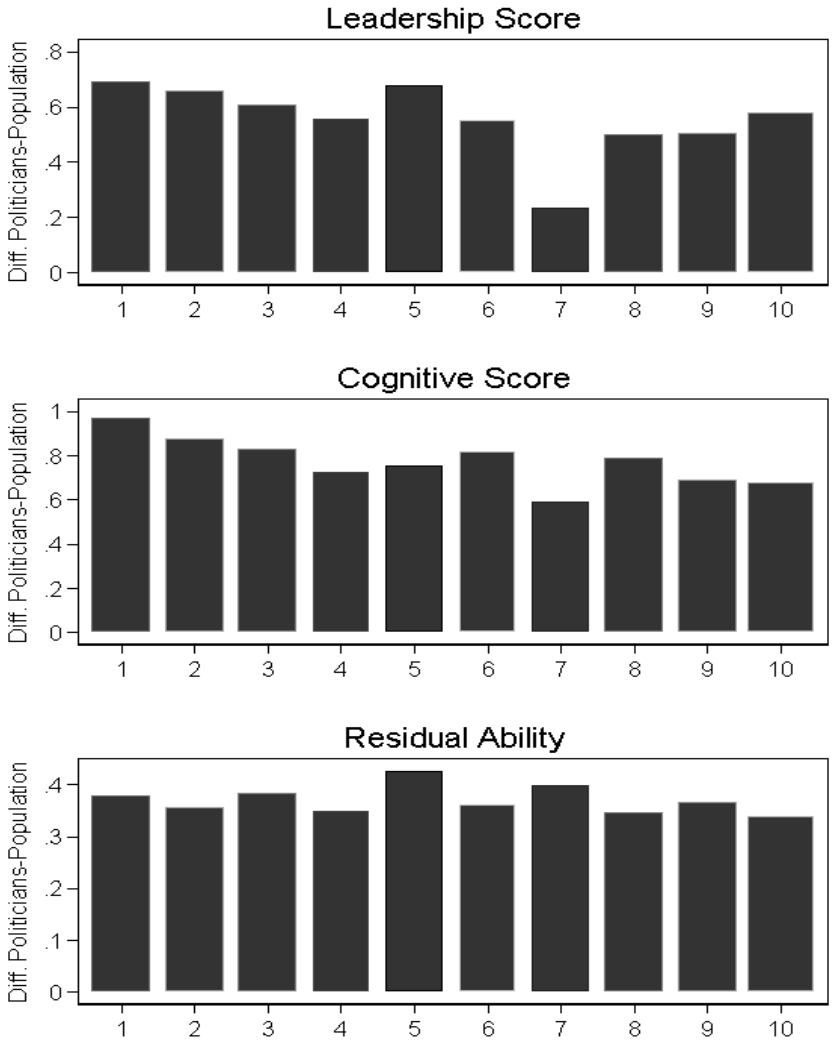

Figure 12: Difference in average ability measures between politicians and general population by father's income decile

Notes: The figure shows the difference in average ability between politicians elected to a municipal council seat and the working-age population, by the politician's father's income decile (x-axis). For each elected politician, or individual in the population, we find their father in 1979 and measure their decile in the working-age population distribution of annual earnings (x-axis). The width of each bar represents the share of politicians whose father belongs to each income decile. The data are pooled for all politicians elected in the 2000s (2002, 2006 and 2010), and the general population is also sampled in these same years. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

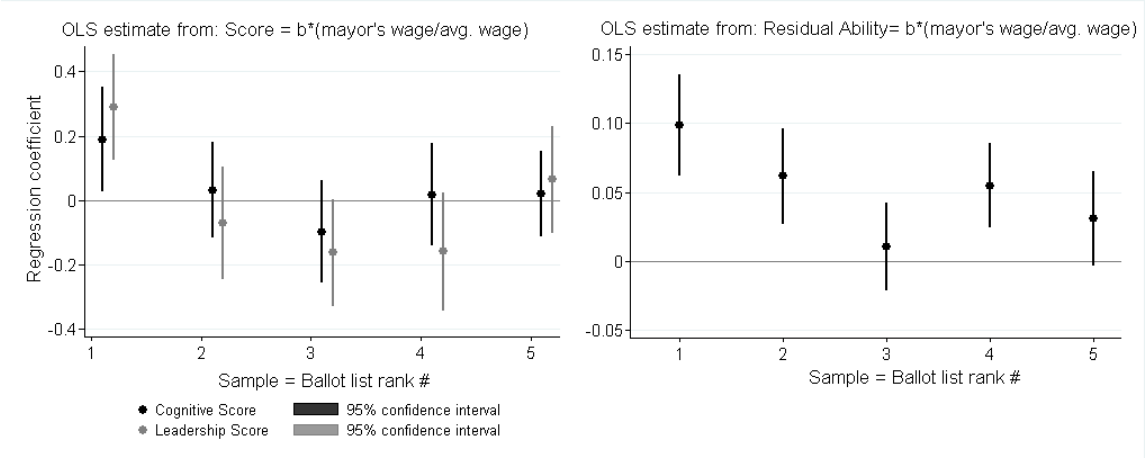

Figure 13: Estimated bivariate relationships between politician ability and the mayor's wage relative to the municipal population, by electoral ballot rank

Notes: The figure shows estimation results for the relationship between politician ability and the mayor's annual earnings as a fraction of the average annual earnings in the municipality that he or she governs. Bivariate OLS regressions are estimated for the sub-samples based on electoral ballot rank. These sub-samples are denoted on the x-axis. In the left graph, the cognitive score (black color) and the leadership score (gray color) are each used as the dependent variable. In the right graph, the dependent variable is the residual ability measure. The dots represent the size of the point estimates and the vertical lines 95% confidence intervals. Data are extracted for the election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006, and 2010). For the enlistment scores, we only use data for men in the 1951-1980 cohorts.

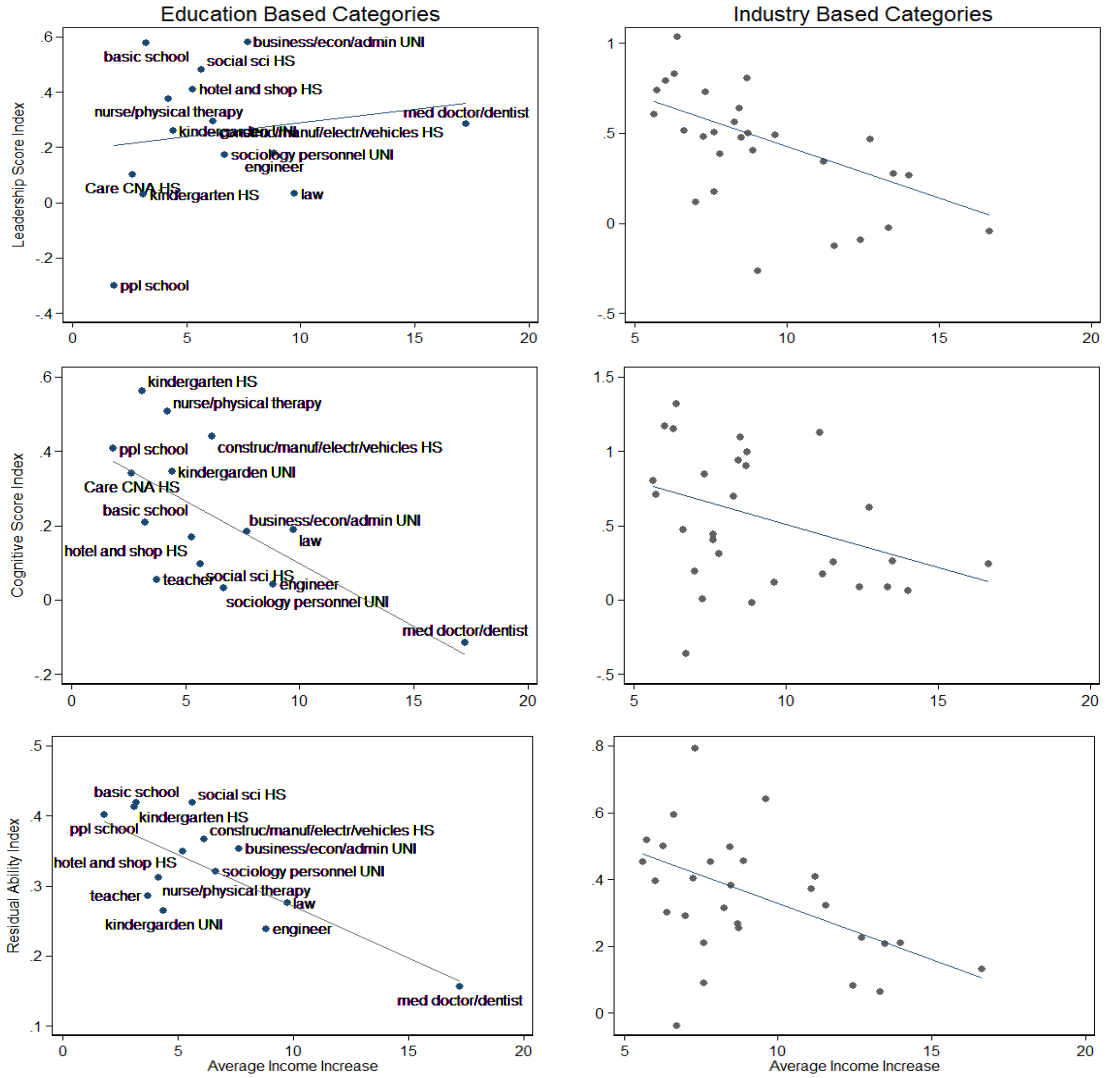

Figure 14: Ability indexes and earnings-tenure profiles

Notes: The figure shows the relationship between the ability selection indices and the income-tenure profile of various occupation types. The x-axis of the left (right) column plots shows the average increase in annual earnings, in 1000 SEK, by education (sector), while the y-axis shows the ability selection indexes by education (sector). Each selection index is computed as mean ability among the elected politicians for each occupation minus mean ability among non-politicians in the same occupation in the working-age population (18-72). Average earnings increases are computed from annual individual-level data for the Swedish working-age population (1990-2012). Individuals with the same occupation are divided into age brackets of five years and year-on-year average earnings hikes are computed for the entire period as earnings(t+1)-earnings(t). The regression lines are estimated by OLS. The cognitive and leadership scores are restricted to 1951-1980 cohort men.

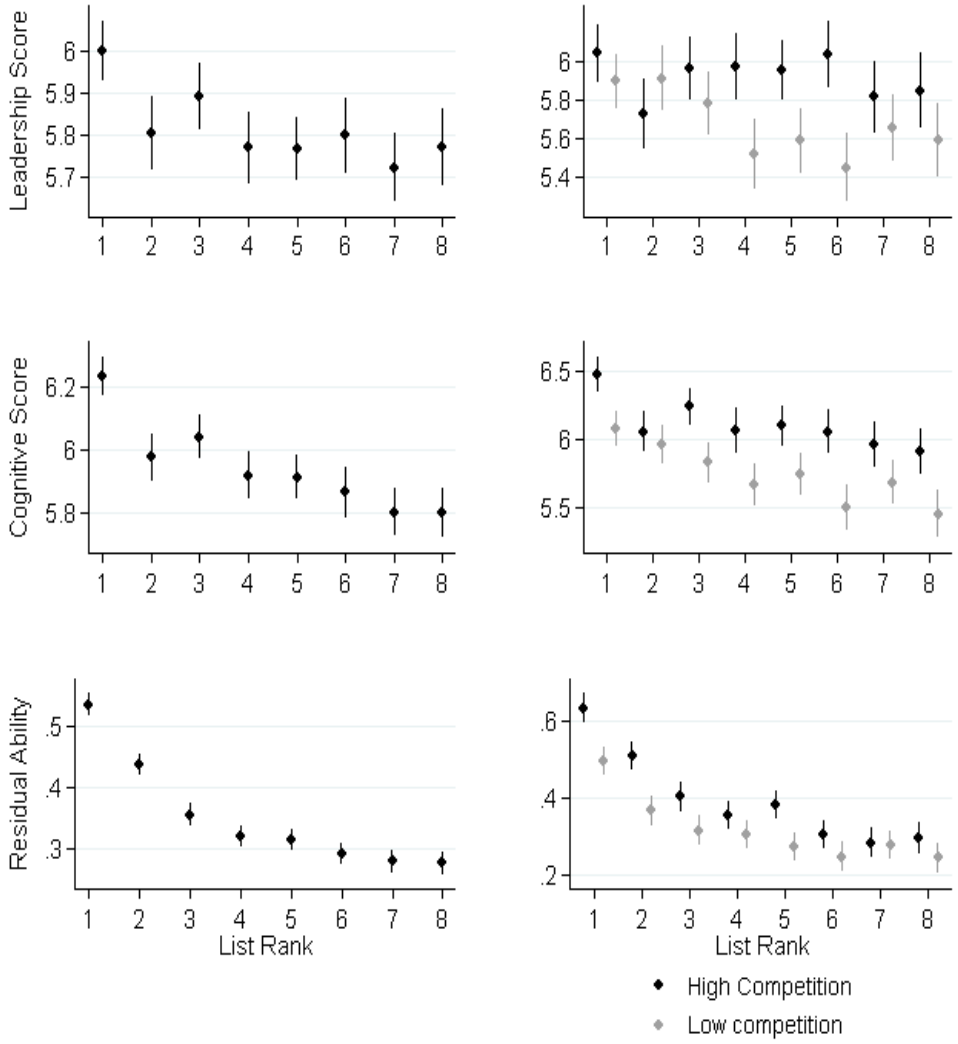

Figure 15: Average ability by party list rank

Notes: The figure shows the averages of three ability variables by ballot rank. The top (middle/bottom) graph shows the means for leadership score (cognitive score/residual ability measure). The data come from elections held during 1991-2010. In the left column of plots, the black dots show the mean of the competence variable for the politicians in each ballot rank, and the vertical lines running through each dot denote 95% confidence intervals. In the right column of plots, similar means and confidence intervals are reported for two groups of municipality-election observations, split by the median of political competition. Political competition is computed as the win margin of the majority bloc of parties, left or center-right. The black dots show the summary statistics for the politicians in high-competition contexts, and the gray dots show the statistics for the politicians in low-competition contexts. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

9 Tables

Table 1: Correlations between ability measures, 2011

	Leadership score	Cognitive score	Years of education
Cognitive score	1		
Leadership score	0.338	1	
Years of education	0.300	0.511	1
Residual ability	0.201	0.167	0.076

Notes: This table reports bivariate correlation coefficients between the various measures of ability. The underlying data encompass the Swedish working age population (18-72) in year 2011. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

Table 2: Summary statistics for the population and politicians, 2011

	Population		Politicians			
	Mean	Std. Dev.	Nominated	Elected	Mayors	MPs
Women	50.5	50.0	41.2	43.1	29.2	44.9
Age	48.8	19.1	53.7	52.4	52.6	48.0
Foreign born (%)	17.1	42.0	9.5	7.6	2.5	8.0
Leadership score (1-9)	5.2	1.7	5.5	5.9	6.4	6.7
Cognitive score (1-9)	5.1	1.9	5.5	5.9	6.3	6.4
Years of education	12.4	3.0	13.4	13.7	13.7	14.8
Residual ability (z-score)	0.02	0.96	0.16	0.46	0.81	0.99
Observations	7,563,148		38,701	12,919	284	349

Notes: This table reports descriptive statistics on social attributes and on the various measures of competence. These statistics were computed for the Swedish working-age population (18-72), and for politicians, both in 2011. Politicians are divided into four hierarchical categories: nominated(but not elected) for a municipal assembly seat, elected for a municipal assembly seat, mayor, and member of (national) parliament. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

Table 3: Ability by selected elite occupations

	Leadership score	Cognitive score	Residual ability	Years of schooling	Labor earnings	Obs
Mayors	6.4	6.2	0.79	13.9	679.4	247
Municipal councilors	5.8	5.9	0.38	13.8	379.0	8870
CEOs ($10-24$ employees)	6.1	5.8	0.81	13.6	675.6	6825
CEOs ($25-249$ employees)	6.4	6.2	1.12	14.2	1046.2	6885
CEOs (≥ 250 employees)	6.8	6.7	1.29	15.4	1926.0	1470
Medical Doctors	6.5	7.4	1.13	17.1	640.0	29514
Lawyers and Judges	6.5	6.8	0.69	17.0	568.0	5308
Economists	5.9	7.0	0.38	20.4	530.0	248
Political Scientists	5.8	6.8	0.61	20.4	513.3	306

Notes: : The table shows ability averages among politicians (mayors in row 1 and elected municipal councilors in row 2) and among individuals in seven occupational categories that make up other "elite" occupations in Swedish society. Columns 1 and 2 reports the means for the two enlistment scores, the leadership score and the cognitive score (measured on a 1-9 scale). Column 3 reports the means of our measure of residual ability (z-score), and column 4 reports average years of education. Column 5 reports the mean of annual labor earnings (in 1000s Swedish Kronor; 1 SEK 0.8USD), and finally, column 6 reports the number of individuals classified into each elite occupation. The data are from 2011 and includes the full Swedish working-age population (18-72). Individuals working in universities are identified based on a 5-digit industry code that indicates employment at a university. For the cognitive score and the leadership score, data is restricted to men in the 1951-1980 cohorts.

Table 4: Average ability selection indices of elected politicians and top-ranked politicians on local parties' electoral ballots, by party-internal career prospects

	Probability that the politician's party appoints the mayor							
	0%		1-50%		51-99%		100%	
	Elected	Top	Elected	Top	Elected	Top	Elected	Top
Leadership Score	0.75	0.95	0.94	1.25	0.79	1.11	0.65	1.52
Cognitive Score	1.29	1.42	1.15	1.37	1.03	1.41	0.9	1.92
Residual Ability	0.35	0.51	0.51	0.79	0.58	0.88	0.6	1.13
Observations*	5,394		1,584		1,032		456	

Notes: The table shows the means of three ability variables by the probability that the politician's party appoints the mayor. We measure this probability as the proportion of election periods that a local party was responsible for appointing the mayor between 1982 and 2010. The means of the leadership score are shown in row 1, the means of the cognitive score in row 2, and the mean of the residual-ability measure in row 3. Row 4 shows the number of local party-election period observations in each category. Data are drawn from the election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006, and 2010).
* For the enlistment measures of ability, only male politicians in the 1951-1980 cohorts are included, which gives a smaller number of observations for these measures.

A Appendix: Supplementary Figures and Tables

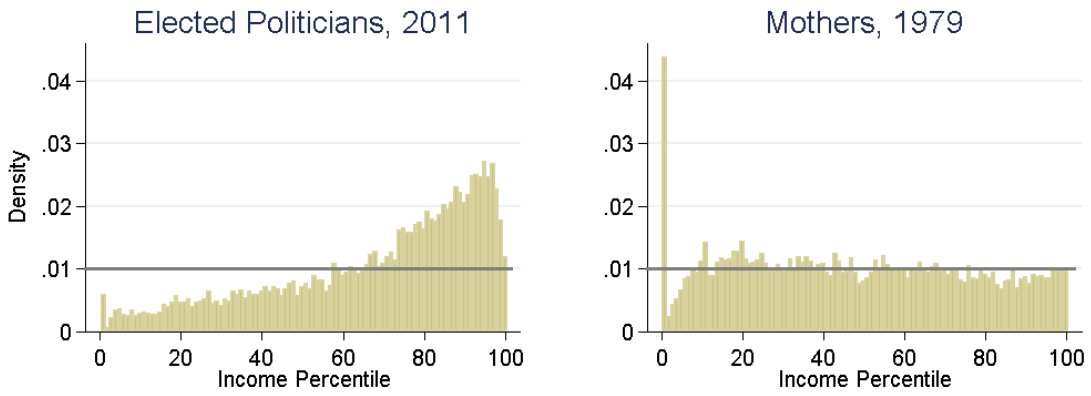

Figure A.1: Distribution of elected politicians and their mothers across the percentiles of population income

Notes: The figure shows the distribution of elected politicians (left) and their mothers (right) across the percentiles of the Swedish income distribution. The income percentiles are calculated by birth year and gender. Data from year 1979 and 2011 were used to compute the percentiles of annual earnings in the working-age population (18-72) in these two years, respectively. The proportion of politicians in each percentile bracket is shown in the histograms. Mothers are included only if they are of working age (18-72) in 1979, and only politicians for whom we can find an earnings observation for their mother are included.

Table A.1: Correspondence between the sub-traits of the leadership score and the Big 5 personality traits.

Sub-trait	Details of component	Corresponding Big 5 trait
Social Maturity	Extraversion	Extraversion
	Having friends	Extraversion
	Taking responsibility	Consciousness
	Independence	Openness
Psychological Energy	Perseverance	Conscientiousness
	Ability to fulfill plans	Conscientiousness
	Ability to remain focused	Conscientiousness
Intensity	Capacity to generate initiative without external pressure	Conscientiousness
	Intensity and frequency of free time activities	Openness
Emotional Stability	Disposition to Anxiety	Neuroticism
	Ability to control and channel nervousness	Neuroticism
	Tolerance of stress	Neuroticism

Notes: : The Swedish enlistment procedure assesses conscripts' psychological traits and combines them into a "Leadership Score." The manual followed by the trained psychologist to make this assessment specifies a process that first scores the conscript on four sub-traits, listed in column 1 of the table, and then merges these assessments into a final score. The sub-traits capture certain aspects of the conscript's personality, which are listed in column 2. In column 3, these personality aspects are related to their corresponding "Big Five" personality characteristics (see Nilsson (2014)).

B Appendix: Formal Proofs

Proof of Proposition 1: The set of willing politicians, the "applicant pool", is $(y,p)\,|p \geq (\pi(1-\delta)+\delta)\,\gamma y - \pi w$ or $y \leq \frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma}$.

The measure of the applicant pool is

$$\int_0^{\bar{P}} \int_0^{\frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma}} g(y)dy\frac{1}{\bar{P}}dp$$

and the average quality in the applicant pool (denoted by $E\,(y_A)$) is,

$$E\,(y_A) = \frac{\int_0^{\bar{P}} \int_0^{\frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma}} yg(y)dy\frac{1}{\bar{P}}dp}{\int_0^{\bar{P}} \int_0^{\frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma}} g(y)dy\frac{1}{\bar{P}}dp}.$$

Denote $\varphi \equiv \frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma} \equiv \frac{p+a}{\psi}$, with $a \equiv \pi w$ and $\psi \equiv (\pi(1-\delta)+\delta)\,\gamma$.

Maximum competence The most competent type willing to enter politics is $\left(\bar{P}, \bar{y} \equiv \frac{\bar{P}+\pi w}{(\pi(1-\delta)+\delta)\gamma}\right)$. Note that $\frac{d\bar{y}}{dw} = \frac{\pi}{(\pi(1-\delta)+\delta)\gamma} > 0$, $\frac{d\bar{y}}{d\gamma} = -\frac{\left(\bar{P}+\pi w\right)}{(\pi(1-\delta)+\delta)\gamma^2} < 0$ showing that the maximum competence increases with w and decreases with γ. Now observe that if $\delta = 1$, $\frac{d\bar{y}}{d\pi} = \frac{w}{\gamma} > 0$ and the statement in the proposition follows from continuity.

Average competence We now prove the statements on average competence.

Effects of w and γ. Since φ is increasing in w and decreasing in γ, establishing $\frac{dE(y_A)}{d\varphi} > 0$ will prove $\frac{dE(y_A)}{dw} > 0$ and $\frac{dE(y_A)}{d\gamma} < 0$. Write out the expression for $\frac{dE(y_A)}{d\varphi}$ as

$$\begin{aligned}
\frac{dE\,(y_A)}{d\varphi} &= \frac{\int_0^{\bar{P}} \varphi g(\varphi)\frac{1}{\bar{P}}dp}{\int_0^{\bar{P}} G(\varphi)\frac{1}{\bar{P}}dp} - \frac{\int_0^{\bar{P}} \int_0^{\varphi} yg(y)dy\frac{1}{\bar{P}}dp}{\left(\int_0^{\bar{P}} G(\varphi)\frac{1}{\bar{P}}dp\right)^2} \int_0^{\bar{P}} g(\varphi)\frac{1}{\bar{P}}dp \\
&= \frac{\int_0^{\bar{P}} g(\varphi)\frac{1}{\bar{P}}dp}{\int_0^{\bar{P}} G(\varphi)\frac{1}{\bar{P}}dp} \cdot \left(\frac{\int_0^{\bar{P}} \varphi g(\varphi)\frac{1}{\bar{P}}dp}{\int_0^{\bar{P}} g(\varphi)\frac{1}{\bar{P}}dp} - \frac{\int_0^{\bar{P}} \int_0^{\varphi} yg(y)dy\frac{1}{\bar{P}}dp}{\int_0^{\bar{P}} G(\varphi)\frac{1}{\bar{P}}dp}\right),
\end{aligned}$$

and use the equality $\varphi = \frac{p+a}{\psi}$, to get

$$= \frac{\int_0^{\bar{P}} g(\varphi)dp}{\int_0^{\bar{P}} G(\varphi)dp} \cdot \left(\frac{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} g(y)dy} - \left(\frac{\int_0^{\bar{P}} \int_0^{\frac{a}{\psi}} yg(y)\frac{1}{P}dydp + \int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} yg(y)dy\frac{1}{P}dp}{\int_0^{\bar{P}} \int_0^{\frac{a}{\psi}} g(y)dy\frac{1}{P}dp + \int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} g(y)dy\frac{1}{P}dp} \right) \right)$$

$$= \frac{\int_0^{\bar{P}} g(\varphi)dp}{\int_0^{\bar{P}} G(\varphi)dp} \cdot \left(\frac{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} g(y)dy} - \left(\begin{array}{c} \frac{\int_0^{\bar{P}} \int_0^{\frac{a}{\psi}} yg(y)\frac{1}{P}dydp}{\int_0^{\bar{P}} \int_0^{\frac{a}{\psi}} g(y)dy\frac{1}{P}dp + \int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} g(y)dy\frac{1}{P}dp} + \\ \frac{\int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} yg(y)dy\frac{1}{P}dp}{\int_0^{P} \int_0^{\frac{a}{\psi}} g(y)dy\frac{1}{P}dp + \int_0^{P} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} g(y)dy\frac{1}{P}dp} \end{array} \right) \right).$$

Changing the order of integration, so that $\int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} yg(y)dy\frac{1}{P}dp = \int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \int_{\psi y-a}^{\bar{P}} yg(y)dpdy\frac{1}{P} = \int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{\bar{P}-(\psi y-a)}{P} yg(y)dy$ and $\int_0^{\bar{P}} \int_{\frac{a}{\psi}}^{\frac{a+p}{\psi}} g(y)dy\frac{1}{P}dp = \int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{\bar{P}-(\psi y-a)}{P} g(y)dy$, we can write

$$= \frac{\int_0^{\bar{P}} g(\varphi)dp}{\int_0^{\bar{P}} G(\varphi)dp} \cdot \left(\frac{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} g(y)dy} - \left(\begin{array}{c} \frac{\int_0^{\frac{a}{\psi}} yg(y)dy}{\int_0^{\frac{a}{\psi}} yg(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{\bar{P}-(\psi y-a)}{P} g(y)dy} + \\ \frac{\int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{P-(\psi y-a)}{P} yg(y)dy}{\int_0^{\frac{a}{\psi}} yg(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{\bar{P}-(\psi y-a)}{P} g(y)dy} \end{array} \right) \right)$$

or

$$= \frac{\int_0^{\bar{P}} g(\varphi)dp}{\int_0^{\bar{P}} G(\varphi)dp} \cdot \left(\frac{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} g(y)dy} - \left(\begin{array}{c} \frac{\int_0^{\frac{a}{\psi}} g(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{(\bar{P}-(\psi y-a))}{P} g(y)dy} \cdot \frac{\int_0^{\frac{a}{\psi}} yg(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy} + \\ \frac{\int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{(\bar{P}-(\psi y-a))}{P} g(y)dy} \cdot \frac{\int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{(\bar{P}-(\psi y-a))}{P} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{(\bar{P}-(\psi y-a))}{P} g(y)dy} \end{array} \right) \right)$$

which can be simplified as,

$$= \frac{\int_0^{\bar{P}} g(\varphi)dp}{\int_0^{\bar{P}} G(\varphi)dp} \cdot \left(\underbrace{\frac{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy}{\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} g(y)dy}}_{A_1} - \left(\alpha_1 \cdot \underbrace{\frac{\int_0^{\frac{a}{\psi}} yg(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy}}_{A_2} + \alpha_2 \cdot \underbrace{\frac{\int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} y\frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}{\int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}}_{A_3} \right) \right)$$

where

$$\alpha_1 = \frac{\int_0^{\frac{a}{\psi}} g(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}$$

$$\alpha_2 = \frac{\int_{\frac{a}{\psi}}^{\frac{a+\bar{P}}{\psi}} \frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}{\int_0^{\frac{a}{\psi}} g(y)dy + \int_{\frac{a}{\psi}}^{\frac{a+P}{\psi}} \frac{(\bar{P}-(\psi y-a))}{\bar{P}} g(y)dy}.$$

Note that $1 \geq \alpha_i \geq 0,\ i = 1, 2$ and $\alpha_1 + \alpha_2 = 1$. Then, if $\int_{\frac{a}{\psi}}^{\frac{\bar{P}+a}{\psi}} yg(y)dy > 0$ (i.e. the support of Y covers a neighborhood of some point on the selection line), $A_1 > A_2$ and $A_1 > A_3$ since $\frac{\bar{P}-(\psi y-a)}{\bar{P}}$ is a decreasing function in y. It follows that $\frac{dE(y_A)}{d\varphi} > 0$.

Effects of π. The effect of π on φ is ambiguous and hence so is the effect of π on $E(y_A)$. This effect becomes unambiguous only as $\delta \to 1$ and y finite. At $\delta = 1$, $\varphi = \frac{p+\pi w}{\gamma}$, and the effect of π is similar to that of w, and $\frac{dE(y_A)}{d\pi} > 0$.■

Proof of Proposition 2: Suppose the party observes types (y, p) perfectly.

Negative screening. Here, the party selects an ability type of zero, which yields more incompetent politicians than the average citizen.

Random (Athenian) screening. This obviously does not require the party to observe types, as it can make random offers to those who volunteer. Clearly, the party obtains politicians with the average competence in the volunteer pool (denoted $E(y_A)$). To show that this is worse than the average competence in the population, we use the identity $\varphi \equiv \frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma}$, noting that $E(y_A) < E(y)$ requires

$$E(y_A) \equiv \frac{\int_0^{\bar{P}} \int_0^{\varphi} yg(y)dy \frac{1}{\bar{P}} dp}{\int_0^{\bar{P}} \int_0^{\varphi} g(y)dy \frac{1}{P} dp} < \frac{\int_0^{\bar{P}} \int_0^{\infty} yg(y)dy \frac{1}{\bar{P}} dp}{\int_0^{\bar{P}} \int_0^{\infty} g(y)dy \frac{1}{P} dp} \equiv E(y).$$

Tis inequality follows from the fact that $\varphi < \infty$ and $E(y_A)$ is increasing in φ.

Positive screening. This is the only form of screening that can deliver politicians who, on average, are more able than the average citizen. By continuity, very mildly positive screening – arbitrarily close to random screening – will yield politicians worse than the population on average. Therefore, positive screening must be pronounced enough for politicians to be better than the average. Under the assumptions that types are perfectly observable and both traits are valuable, the party will select types $\left(\bar{P}, \frac{\bar{P}+w\pi}{(\pi(1-\delta)+\delta)\gamma}\right)$.■