Who Becomes a Politician? *

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

February 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 participants in seminars at CIFAR, UDESA, Berkeley, Bocconi, LSE and LBS for helpful comments and to the Söderberg Foundation and the Swedish Research Council for financial support.

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. As a collective, voters may 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 free-riding incentives (Olson 1965; Messner and Polborn 2004) and lower opportunity costs (Caselli and Morelli 2004). In addition, if ability is distributed unequally in society, 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 candidate screening by voters and/or parties. The study of candidate screening requires 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 data limitation to advance our understanding of candidate selection.[3] Unfortunately this literature misses candidate entry, which requires information on those who did not even attempt political entry.

Second, much of the relevant theory emphasizes the quality of political selection.

[1] See for example 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).

But how to measure quality or competence? Absent direct data on the underlying intelligence or personality traits of politicians, the existing empirical literature has relied on proxies such as education, occupation, or pre-office income levels (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). While correlated with competence, these proxies also likely reflect luck or social class.

Third, representation is hard to measure. Previous work has relied on proxies like a politician's 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, 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 detailed administrative data from Sweden. The focus on Sweden deserves some justification beyond data availability. Clearly, conclusions based upon evidence from a single country cannot be extrapolated to the rest of the world. But Sweden is of interest for its commonalities with other countries, as well as for its differences. Swedish politics is based on an electoral party 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 (scoring a perfect 10 in the -10 to 10 Polity-IV scale) for a long period, during which other states experienced a gradual transition from autocracy to democracy. Sweden thus exemplifies the 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, defenders 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 political selection in a country. A key advantage is the availability of rich measures of competence and social background. The competence measures include evaluations of IQ and leadership potential done by the military on the universe of the male population, and a measure of residual ability (for both genders) stemming from a fully saturated Mincer regression that was developed by Besley, Folke, Persson, and Rickne (2013). Our social-background data includes

reliable intergenerational information, namely social class as reflected in parental occupations (a traditional approach in sociology) as well as parental income.

We focus on municipal politicians – who are important providers of public goods – to take advantage of larger numbers and within-country variation. We approach these data at three different levels of analysis that make for the three main sections of the paper.

Aggregate level At the highest aggregate 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 prospective politicians have higher pre-office incomes and more promising careers and hence face higher opportunity costs from entering public life. Nevertheless, we document strong positive selection along all competence measures. Selection monotonically improves from those merely nominated to those elected, and from 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 individuals, the distributions of their parents' social class as well as parental earnings look almost identical to those of the population. In other words, politics is not reserved for the scions of rich elite families but tracks the entire social structure. This remarkably even representation is the result of different political parties representing different segments of the income distribution.

The findings on representation help us understand the drivers of positive 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 human-capital (itself manifested in competence measures) *cum* easier access to political power. 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 findings on representation indicate that family background is not an important driver. Moreover, when we compare the competence of politicians and their siblings, we find differences nearly as large as those separating politicians from the general population. This indicates that, conditional on family background, individual traits matter greatly for selection.

Municipal level In the second part, we examine municipal heterogeneity. Here, we test whether municipalities that display stronger selection 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 competence among lower socioeconomic strata.

Individual level Finally we consider the decisions of individuals about entry into politics, and the role that political parties may play in allowing them access to political office. A highly stylized formal model, which shares some features with the aforementioned economic selection models, guides our thinking on drivers of self-selection. It predicts that (i) higher mayoral wages attract more competent individuals not only to the mayor position, but also to engage in politics. The model also predicts selection into politics to be improved by (ii) intrinsic motivation 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 predictions.

Our analysis also indicates that, given self-selection choices, the strong positive selection on competence likely reflects positive screening by political parties. Parties may serve as arenas where the more able are allowed in, and promoted to positions of influence. This hypothesis matches qualitative work in political science, which describes the active involvement of parties in promoting the more competent. It also matches two facts in our data, namely that competence increases with list ranks within parties, and that competence for each list rank increases with levels of political competition within municipalities.

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 tradeoff between selecting on ability and achieving broad representation. 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 and turnout is 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 focuses on the latter.

In Sweden's proportional representation system, seat shares in the municipal councils are proportional to the vote share of each party. Until 1998, seats for each party were handed out from the top of the ballot (so-called closed lists). 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 held representation, but these parties are rarely part of governing majorities. Parties that run 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 government – the members of which are appointed by the governing majority to mirror the seat distribution. The largest party within 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 importance of the mayor partly reflects that municipalities play a crucial role in the economy: municipalities carry out expenditures which represent about 25 percent of Sweden's GDP and employ 20 percent of the country's workforce. Municipal politicians are thus responsible for public-goods provision in 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. Electoral competition can be illustrated by the changing proportion of left-bloc governments over time. This proportion increased from 31 percent in 1991 to

[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.

73 percent in 1998, only to fall back down to 59 percent in 2002 and 41 percent in 2006.

Running for office and political careers Depending on the size of the municipality (from 2,558 to 780,817 inhabitants), members of the local parties are divided into "clubs", formed by neighborhood associations and other social groups, such as the women's club and youth clubs. A citizen who enters politics becomes member of a party and joins one or several clubs. There are no routes to office outside of existing parties (except that individuals can form new parties). A political career may involve moving to the top of the local-party hierarchy – in the case of the largest majority or minority parties, this also implies a position as mayor or vice mayor. A small fraction of politicians move on to a parliamentary seat, but municipal politics is a breeding ground for national level politicians. Among parliamentarians elected in 2010, 72 percent had been elected to a municipal council at some point after 1982.

Local parties compose their own electoral ballots without interference of the central party organization. In the two left-bloc parties (Social Democrats and Left Party), party clubs nominate members to a committee that aggregates the nominations into a proposed list. In the other parties, the nomination committee organizes internal test elections, where few members exercise their right to vote. At the end of the day, the leadership of the party dominates the nomination committees and largely controls the composition of the ballot.

Sweden has a strong normative tradition of citizen politicians (so-called "leisure politicians"), whereby political appointments are regarded as spare-time activities that complement one's work in the regular labor market. Consequently, almost all elected political officials receive only piece-rate compensation for time spent in meetings. Previous research has also shown that there are no indirect monetary gains to winning a seat in a municipal assembly (Lundqvist 2013). However, the mayor is a full-time employee, and in most municipalities the vice mayor also gets a part-time salary. The mayor's wage, determined by the municipal council, is typically in the top percentile of the Swedish earnings distribution. Moreover, becoming a mayor is associated with high status and substantial power.

The monetary costs to running for municipal office are minimal. All politicians run as part of a party ballot, and campaign finances are 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).

A qualitative literature on local Swedish politics suggests that the motivation of

citizens to enter politics is connected to intrinsic concerns with policy, as well as a desire for social interaction in policy circles (Karlsson 2001). However, pecuniary concerns are also present, especially as experienced leisure politicians contemplate more serious appointments (Dahl 2011).

Voters preferences and selection Swedish citizens value a political class that is both competent and socially representative. When asked about what motivates their party choice, the competence of politicians has remained one of the top three reasons in the past decades, over time rising to the same level of 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 and social groups, as well as geographical areas (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 classified 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 the basic patterns of 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 the sample definitions used in our analysis.

Sources Our empirical analysis is based on data from three types of sources. The first dataset contains a list of elected and non-elected individual candidates that ran for political office during the period of 1982-2010, approximately 200,000 unique individuals. Prior to an election, each political party must report its ordered ballot or list, which contains 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 comprises approximately 50,000 elected individuals.

In a second step, we link these data on politicians to several administrative registers held by Statistics Sweden. These databases contain annual data for the whole Swedish population ages 16 and above. For most variables, we have access to records for the entire population on an annual basis from 1979-2012, 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 are able to precisely characterize how the personal attributes of politicians relate to the distribution 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 because 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 year 2003. To bridge this gap for earlier periods, we further complement the occupation data with information from the Population Censuses available every five years.

Our final individual-level dataset comes from the Swedish Defense Recruitment Agency. For a subset of cohorts, the 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 the cohorts of 1951 and onward. For quality reasons, we also truncate the data for men born after 1979. For cohorts born between 1951 and 1980, enlistment rates for men were around 90 percent.

Electoral results are linked to our dataset at the municipal level by using records kept by the Swedish Electoral Agency. This gives us the proportion of votes received by 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 and mental 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 of the sample are younger than 18 or older than 19). The timing of enlistment generally corresponds to the year that the person graduates from high school, so the test scores are not influenced by university training. Because these tests tended to be high-stakes – better results generally resulted in more interesting and rewarding military placements – the data quality is considered high.

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 the results from a 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 Armed Forces Qualification Test (AFQT) in the United States, which focuses more on "crystallized" intelligence, i.e., skills that are teachable. 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 For individuals who score a 5 or higher on the cognitive test, a more in-depth evaluation is done for possible placement into a position of military leadership. 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 ability (from the test described above), 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 it is known that the test is designed to evaluate a conscript's personality in civilian life, as well as 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

[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.

related to the well-known big-5 personality traits (extraversion, consciousness, openness, conscientiousness, and neuroticism) – see the Appendix, Table A.1.

Education Educational attainment of all individuals is reported by Swedish schools and universities to the relevant authorities, and records are kept by Statistics Sweden. For people immigrating into Sweden later in life, information on schooling is collected through surveys. Education levels are categorized according to a Swedish standard. These 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 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) who use the 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 13 one-digit industrial codes.[6] The function f represents

[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 observation 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".

a full 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]

Is residual ability a good measure of competence for politics rather than merely 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 a number of measures of political success.

Summary statistics Table 1 displays pair-wise correlations between our main competence measures for the Swedish male population. (We will use the terms competence and ability interchangeably throughout the paper.) While all measures are positively correlated, they are clearly measuring different ability traits. 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 with a correlation coefficient of 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]

[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 parental leave and engage in care activities that drive an increase in the gender pay gap when couples have children.

As for retirement, there are plenty of senior politicians. Mincer equations of the retired and people in the workforce differ as retirees do not have a current employment sector. We compute the income residuals of retirees based on their sector of employment in the majority of their working-life. For people who retire during the sample period, we only include the years when they were in the workforce to calculate their residual ability.

Table 2 reports means of the ability variables for politicians and the entire population (both subject to availability) in the year of 2011. For municipalities, we distinguish between nominated (but not elected) and elected politicians, and mayors. For comparison's 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. But the most striking point of the data is the subject of 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 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. As we saw, two measures – education level and residual ability – are available for the entire population, while the other two – cognitive and leadership scores – are available only for males.

Cognitive and leadership scores The top-left panel in Figure 1 shows overlapped histograms for the cognitive 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 cognitive 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 5.1 are more highly represented among nominated politicians than in the general population.[8] For elected

[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

politicians, the shift to the right is stronger, and even more so for mayors. Mayors have more than a full additional point (out of nine) higher cognitive scores than the population, with an average score almost 2/3 of a standard deviation higher.

[Figure 1 about here]

The top-right panel shows a similar graph for the leadership score. Politicians score higher than the average Swede, more strongly so when elected to office, and particularly so when selected for the top-municipal office. Mayors score an additional 1.2 points higher than the average person (who took the test), which amounts to 70 percent 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 for the population. The elected display a clearer shift, with a 40 percent of a standard-deviation difference. Mayors have the highest residual ability levels, with a mean 2/3 of a standard deviation above the population mean.

The evidence on residual ability is important for at least two reasons. First, it includes both females and males. Second, strong positive selection on intelligence and leadership alone might just reflect that those who become politicians face lower opportunity costs. 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 an extra year of education than the average Swede, whereas elected politicians and mayors have 1.3 years extra. In contrast to the other ability measures, the politician distribution of educational attainment does not first-order stochastically dominate the population distribution: politicians are underrepresented at the highest level of education, which includes PhDs. In other words, politicians are positively selected in terms of education levels and years, but doctoral–type degrees are less prevalent among politicians than in the population. In the remainder of the paper, we focus on the non-education measures of competence.

The key takeaway from these graphs is that a strong pattern of positive selection appears to dominate Swedish politics and that selection gets more positive the higher

otherwise.

a person's position in the political hierarchy. This occurs despite the fact that more competent people face higher opportunity costs from entering public life.

Politicians vs. other high-status professions To gain an additional perspective on selection, we also compare politicians to individuals in some elite occupations in Sweden, 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 mayoral earnings are substantially lower. 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, go on to play leadership roles.

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 the scions of the richest 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]

[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.

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

Politicians and their siblings Figure 2 compares the distribution of competence traits of elected politicians with that of their siblings. It shows that 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. This comparison shows that the difference vis-a-vis siblings in terms of cognitive scores can account for 73 percent of the gap vis-a-vis the population. For leadership scores the corresponding number is 80 percent, and for residual ability it is 63 percent. This evidence is a strong indication that competence, rather than family background, is the key selection criterion.

[Figure 2 about here]

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.

Figure 3 gives additional evidence that background is less important than own characteristics. It classifies politicians according to their percentile in the income distribution of the entire population. By definition, the general population would display a perfect uniform distribution with a density of 0.01 for each percentile. 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 competence traits rather than to family characteristics, since the latter would naturally extend to politician siblings.

[Figure 3 about here]

Politicians and their parents We next examine the relevance of social background, and show that it does not matter much, in the aggregate, for political

selection. We characterize social background through parental income and occupational status, and show that politicians do not come disproportionately from elite backgrounds.

Specifically, 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 of any age, but the results remain the same when we instead restrict the analysis to fathers 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 the percentiles of the annual-earnings distribution. 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 persons with high earnings relative to the population. But for fathers, the distribution has a uniform shape: in terms of parental earnings, 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 medical doctors (the upper panel) and CEOs of firms with 10-25 employees (the lower panel). As expected, the 2011 earnings for doctors and CEOs (in the left graph of each panel) disproportionately fall into to the upper percentiles of the population income distribution, even more so than for politicians. However, the corresponding figures for the fathers of these doctors and CEOs (in the right graph of each panel), show that their 1979 earnings are also much more skewed to the right than the earnings of the politician fathers in Figure 4.

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 terms, as well as in relative terms (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 left graphs of these panels 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. Among the politicians' parents in the right graphs of the panels, however, we clearly see representation of different levels of earnings. High-earners are over-represented among fathers of Conservative politicians, while middle-income parents are over-represented among Social-Democrat fathers. Low-income earners are over-represented among Center-party fathers, who 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 combination of diverging party representations renders the almost perfect representation of parental incomes displayed in Figure 4 for the aggregate of politicians. Of course, this is not a coincidence but an illustration of the presumption that different parties will represent different interests – at least in societies with a multi-party system where the left-to-right dimension plays an important role in politics.

Social class rather than income Although parental earnings are informative of social representation, they may only capture a part of the social structure. Therefore, we supplement the information on parental income with information on parental social class whenever the relevant information is available. Figure 7 compares the distributions of social class for politician parents and the population. The class division is based on three variables: occupation, education level, and gender, and corresponds closely to the EGP social-class scheme (Erikson and Goldthorpe 1992). 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.[10] The data are again from 1979, and contain a match for at least one parent for 69 percent of the politicians. This time, we do not limit ourselves to fathers but use the parent with the highest social class.

[Figure 7 about here]

[10]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.

The pattern in Figure 7 corroborates our previous findings 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 would be that meritocracy and elitism are not rival explanations, but one and the same. In a world where individual competence is helped by parental investments in human capital, a strictly meritocratic system will favor elites. Meritocracy would then favor the competent within a family, but still be elitist across families. However, the very even representation of different social classes that we find rejects this interpretation. Our preferred interpretation is that Sweden's political system, on average, is both meritocratic and broadly representative.

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 selection and representation at the municipal level, we compare the politicians in each municipal council to the general population in the municipality. We construct a simple selection index, as follows. For a competence 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 of index to gauge ability (when k is a competence score), as well as representation (when k is the ordered value of a variable 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 by definition has a uniform distribution with an average percentile of 50. If politicians are positively selected on competence in that municipality, they are

drawn on average from percentiles higher than 50, so the selection index will be positive. In municipalities with negative selection, politicians will be disproportionately drawn from percentiles below 50, and the index will be 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.

Distribution of competence across municipalities Figure 8 shows the distribution of the political selection index for the three ability measures using municipality-level data for elections during the 1991-2010 period. As indicated before, 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. But despite this positive selection of competence on average, we see considerable spatial variation with a non-trivial share of municipalities exhibiting adverse selection. Based on our residual ability index, 4.4 percent of municipalities selected mayors below their population average (77 out of 1,733 municipality-election observations).

[Figure 8 about here]

Distribution of representation across municipalities Figure 9 shows plots of 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 index for parental income (the right plot) is more or less centered at zero, as we would expect from Figure 4.

The distribution of the selection index for parental social class (the left plot) has mostly positive support, with modal values between 0 and 0.5. The key to this is the pattern in Figure 6, which shows that politicians with farmer parents (occupation value 4) are over-represented (through the Center party). This is more pronounced in smaller municipalities that have smaller councils (as the Center party is strongest in those councils). Because the plot in Figure 9 compares municipalities while Figure 6 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 when it comes to both competence and representation. It is natural to

ask whether the two aspects of selection display positive covariance. Suppose that a certain municipality tried to improve its selection along the competence dimension by recruiting politicians from higher socioeconomic background. Then, positive ability selection might come at the cost of worse representation. If this were generally true, we would expect a positive correlation between municipal competence and representation indices.

In Figure 10, we plot binned averages of our two representation indexes (for parental income, and parental social class) against each of three competence indices (for cognitive score, leadership score, and residual ability). Together with the figures, we also provide the bivariate correlation for each relationship (b) and the normalized relationship in terms of standard deviations (*beta*). Positive correlations suggest positive relationships between an overrepresentation of parents of higher earnings, or upper social classes, and ability selection.

[Figure 10 about here]

Overall, these plots suggest a very weak tradeoff between competence and social class (left column of plots). The estimated slopes are all positive, but the slope coefficients are small. The strongest relationship, between parental social class and politicians' cognitive scores, suggests that a one standard deviation higher over-representation of higher social classes is associated with a 0.13 standard deviation higher cognitive score. For the two other competence measures, the relationships are much weaker.

Turning to the relations between competence and over-representation of higher parental incomes (right column of plots), we find the strongest relationship for residual ability. A 10 percentile increase in average over-representation of parental income is associated with a higher average residual ability of 0.03 (with a corresponding *beta* estimate of 0.12). 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 there only a tenuous relationship between social representation and competence? The flat tradeoff may be particularly surprising, given the measurement problems in disentangling innate ability from parental background. After all, enlistment concerns 18-year olds, and even though Sweden has a comparatively uniform education system, parental background is likely to affect outcomes 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 competence varies across parental socioeconomic groups.

[Figure 11 about here]

Figure 11 plots our three measures of competence by parental social class (occupation) for politicians, as well as for the general population. Despite the broad access to education in Sweden, the cognitive and leadership scores do correlate positively with parental occupation. Our residual ability measure does not, probably because it reflects the residuals from a highly saturated Mincer equation.

In the figure, the number above each pair of bars shows the (absolute) difference between politicians and the population. As these numbers indicate, positive selection on competence tends to be larger in lower social classes, which mitigates the competence cost of recruiting politicians from less favorable backgrounds. This is particularly the case for the cognitive score, where we clearly see that politicians from the lowest parental social class are more positively selected than those from the highest. These patterns help explain the lack of any substantial tradeoff between competence and representation documented in Figure 10. Figure 12 shows similar patterns for politician competence by deciles of parental income.

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

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 pre-existing party members 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

In this subsection, we write down and analyze a simple "Roy model" of self-selection into politics, and consider 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: Y

(with typical element y) a measure of ability and P (with typical element p), a measure of the intrinsic motives to serve (PSM, for public service motivation). To simplify the comparative statics, we assume that both parameters are non-negative, bounded above, and jointly uniformly distributed over a convex set: $(y, p) \in T$ with $y \in \left[0, \bar{Y} > 0\right], p \in \left[0, \bar{P} > 0\right]$.[11]

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] Someone who 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 (which can vary across professions).[13]

Someone who offers to serve in politics gets accepted to run and is elected council member with probability $q(y)$ – we consider different slopes of the screening $q(y)$ function below. Elected politicians get non-pecuniary benefits $\frac{p}{2}$ in each of the two periods. They also have to give up some career: their first-period private earnings are y, but second-period expected earnings are $(1-\delta)\gamma y$. In other words, the opportunity cost of politics is $\delta\gamma y$, with $\delta \leq 1$ – it depends not only on general ability, but also on private-career prospects.

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

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

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

That is, the intrinsic return to politics (the first term of the LHS) plus the probability of an income gain when becoming mayor (the second term on the LHS) has to outweigh the opportunity cost of giving up some career prospects (the RHS).

[11] We can also assume that y is independent from the uniformly distributed p, with an arbitrary (continuous) distribution.

[12] At the cost of introducing complexity, the model can be extended to include discounting as well as sequential decisions: a person who entered into politics 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 δ 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.

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 this line want to enter and those below want to 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 out and the set of those willing to enter gets larger. Given our assumption of a uniform distribution, it is straightforward to show that average ability must go up and average PSM must go down as the mayoral wage rises.

For parameter π we get,

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

which in general is ambiguous in sign. This is because the selection line pivots. For ability types close to zero, the derivative approaches $-w$, meaning that the line shifts down. But for very high-ability types, the derivative is positive if δ is low enough, and the selection line shifts up. If parties recruit from the high-ability segment, this would mean that a higher appointment probability reduces the supply of high-ability types. As we shall see, however, the data tell a different story, which can only be the case if δ is high enough so that the line shifts down for the entire type space (for any finite $\overline{Y}$, δ close enough to 1 makes $\frac{dp(y)}{d\pi}$ unambiguously negative).

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

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

meaning that higher future income in the labor market shifts the line up. This discourages entry, lowers average ability, and raises average PSM. A steeper earnings-tenure profile thus makes for worse supply.

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

Proposition 1 *If (p, y) are drawn from a joint uniform distribution, maximum as well as average competence of people self-selecting into politics increase (weakly) with (a) higher w and lower γ, and with (b) 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 that were 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 the role of intrinsic motives and dynamic career returns.

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 $q'(y) < 0$); and (iii) positive selection (i.e., meritocracy, with $q'(y) > 0$).

Since Sweden is a party-based democracy, we adopt the fiction of a party planner who selects from the available pool of candidates, anticipating voter demands. The main question is whether our earlier finding that elected politicians have higher ability than the average citizen necessarily 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.*

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 in order 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.

Consider first random selection. Given the selection line $p(y) = \pi((1-\delta)\gamma y - w) + \delta\gamma y$, the earlier comparative statics show that expected candidate ability (denote it $E(y_A)$) must be worse than the average ability $E(y)$ in the population. To see this, (and abstracting from the fact that $\gamma \geq 1$) it is convenient to note that the entry condition implies that $p(y) \to \pi w$ 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, $E(y_A)$ must lie below $E(y)$. Given this, the result for negative screening is obvious, because the average candidate must have ability even worse than $E(y_A)$. The argument with imperfect screening is less direct, but also feasible.

Positive screening implies that we can use our comparative statics of supply to characterize those who are selected into parties and elected. This is easy to see if the party observes types (p, y) perfectly.[14] If the party does not value motivation but values competence, then it only selects individuals on the line $p(y)$, and any change in (γ, π, w) that shifts the line down will increase the average and top quality not just among those willing to enter, but also among those elected. If the party values both motivation and competence it will select individuals of type $\left(\bar{P}, \min\left\{\frac{\bar{P}+w\pi}{(\pi(1-\delta)+\delta)\gamma}, \bar{Y}\right\}\right)$. Thus, the competence of politicians is weakly increasing in w and π (when δ is high enough), and weakly decreasing in γ.

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

To relate the value of a certain wage to income opportunities in the municipality, we normalize the mayor's annual earnings by the average earnings among all municipal 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 have ever appointed a mayor in the period 1991-2010. Because the probability of becoming mayor varies by rank on the municipal party list – top-of-the-list politicians being the most likely, the

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

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 a 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, which captures the relationship between the mayor's relative wage and the selection of politicians for a particular list rank. If a higher mayoral salary attracts higher quality individuals, β_r should be positive. Moreover, if high earnings attracts high-ability individuals to seek positions with a higher probability of becoming mayor, β_r should be higher for $r = 1$.

Figure 13 plots our estimate of interest from five versions of equation (4), one for each r-sample. The left graph shows the estimates for the cognitive and leadership scores. There is a positive and significant relationship only between the mayor's wage and the selection into top rank. For the remaining ranks, the estimates are smaller in size, close to zero, and not statistically significant. The right graph shows the estimates for residual ability. Here, we see 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 it highly probable that they will appoint the mayor. Other parties are small in size and belong to the political opposition, making it highly unlikely for them to appoint the mayor. We now use the electoral support for parties and blocs to classify parties into categories by the political career opportunity they afford.

To do so, we select a sample of stable political conditions: we only consider municipalities where one bloc, left or center-right, has held a majority in five consecutive election periods. This sample entails 47 percent of all municipality-election observations. To further differentiate career opportunities between large and small

parties within each bloc, we drop blocs where any smaller bloc member has a vote share within 75 percent of the largest bloc member in any of the three most recent elections. Finally, we exclude parties with less than three seats, as a "career" is less meaningful in a party with one or two elected politicians. Taken together, these restrictions leave 27 percent of the original sample.

Within this sample, we divide parties into four groups: (1) largest party in the governing bloc, (2) smaller parties in the governing bloc, (3) largest party in the opposition bloc, and (4) smaller parties in the opposition bloc. From appointment data for the 2006 and 2010 elections, we can verify that our categorization does produce substantial variation in this proxy for π.[15]

Next, we compare competence selection indices across these four categories. We want to know if: (i) parties of type (1), namely the largest party in the governing majority, stand out in terms of positive selection, (ii) parties of type (2)-(4) with a near-zero probability of appointing the mayor 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. As for point (i), we find that dominant majority parties indeed have a better selection of their top-ranked politician.[16] 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 As for point (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 category with the weakest career prospects as in the party categories with better prospects.

In terms of our model, this suggests that economic motives for self-selection tells only part of the story. In addition, intrinsic motivation must play a role beyond the material rewards of municipal office. 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 $p > 0$.

[15]Table A.2 in the Appendix shows the proportion of parties in each category that appointed the mayor, any full time position (recall that the mayor is not always full time), any full time or part time position, or a parliamentarian.

[16]Majority parties have larger party delegations on average, which means that average competence among the rank and file is pulled down by moving further down the competence distribution to elect a larger number of politicians.

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 first compute earnings-tenure profiles for different occupations. We do so 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 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 for politicians who belong to each of the different labor-market segments. 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 in the left column and employment sectors (industries) in the right column.

[Figure 14 about here]

As is evident from the figure, the data are consistent with the prediction from our simple model. All the graphs, except one, 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 screening of candidates. This subsection provides some evidence for such positive screening.

Background Parties can screen candidates in different ways. One way is to provide an arena where politicians can compete with one another in coming up with the best arguments and policy proposals. Such competition may well result in positive selection if more able politicians win out in this tournament-like environment and as a result 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 a competence 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 competence levels vary by list rank for the cognitive score, the leadership score, and residual ability, respectively. 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 a significantly higher ability than every other rank for all three competence measures.

[Figure 15 about here]

We thus have some evidence that parties indeed screen and promote more competent 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 – those that can offer positive screening to society – thus appear as an important component of democracy

Selection and political competition If political parties engage in screening and also want to win elections, then parties operating under stiffer political competition should have stronger incentives to screen their candidates for competence. To see whether 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 both theoretical arguments and empirical evidence for the notion that leaders matter, and that societies benefit from competent and broadly representative leadership. While democracy may be well suited – relative to other political systems – to promote representation, it is not clear that democracy can deliver leadership that is both representative *and* competent. In this paper, we analyze selection into politics with regard to competence and representation in Sweden, a paradigmatic advanced democracy. We use a unique data set, with rich information on competence traits and social background for the universe of municipal politicians and for the entire Swedish population. Based on these data, we uncover four new 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 competence and representation; (4) Individual motives seem to 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 indicate that representative democracy, at least in one of its advanced incarnations, is compatible with highly competent and representative leadership. This is important because it alleviates a concern that political systems that encourage 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 have important implications for how to think about political selection and recruitment. Standard models of candidate entry, which focus mainly on the extrinsic motives of holding office and on opportunity costs, generally cannot explain positive selection. The data seem to support a richer model, in which 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 competence is an important element in this calculus. But what criteria are used and how the process actually works in the internal party organization remain open questions in need of further empirical and theoretical work.

Future research should also extend the analysis in several dimensions. There is room for more comprehensive modelling to interpret and better guide the empirical work, especially when it comes to the roles for parties and voters. In addition, future research should attempt to quantify intrinsic motivation, and its impacts on selection and performance. Finally, it would be valuable to add to the evidence on the relevance of leaders, by studying whether and how the competence and representativeness traits we study 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., 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 folkrrelsepartiet: Om aktivism och politiska kursomlggningar 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 jmstlldhet? (Who wants gender equality?). In S. Holmberg and L. Weibull (Eds.), *Det vras fr politiken:*

trettiotv artiklar om politik, medier och samhlle (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.

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.

Freidenvall, L. (2006). Vgen 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.

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 frtroendevald i kommuner och landsting (New to local politics). In *Att vara med p riktigt demokratiutveckling i kommuner och landsting (To participate for real - democratic developments in municipalities and counties).* SOU 2001:48, Swedish Government White Paper.

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.

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

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge: Harvard University Press.

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.

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, Allmnna valen Valunderskningen (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.

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

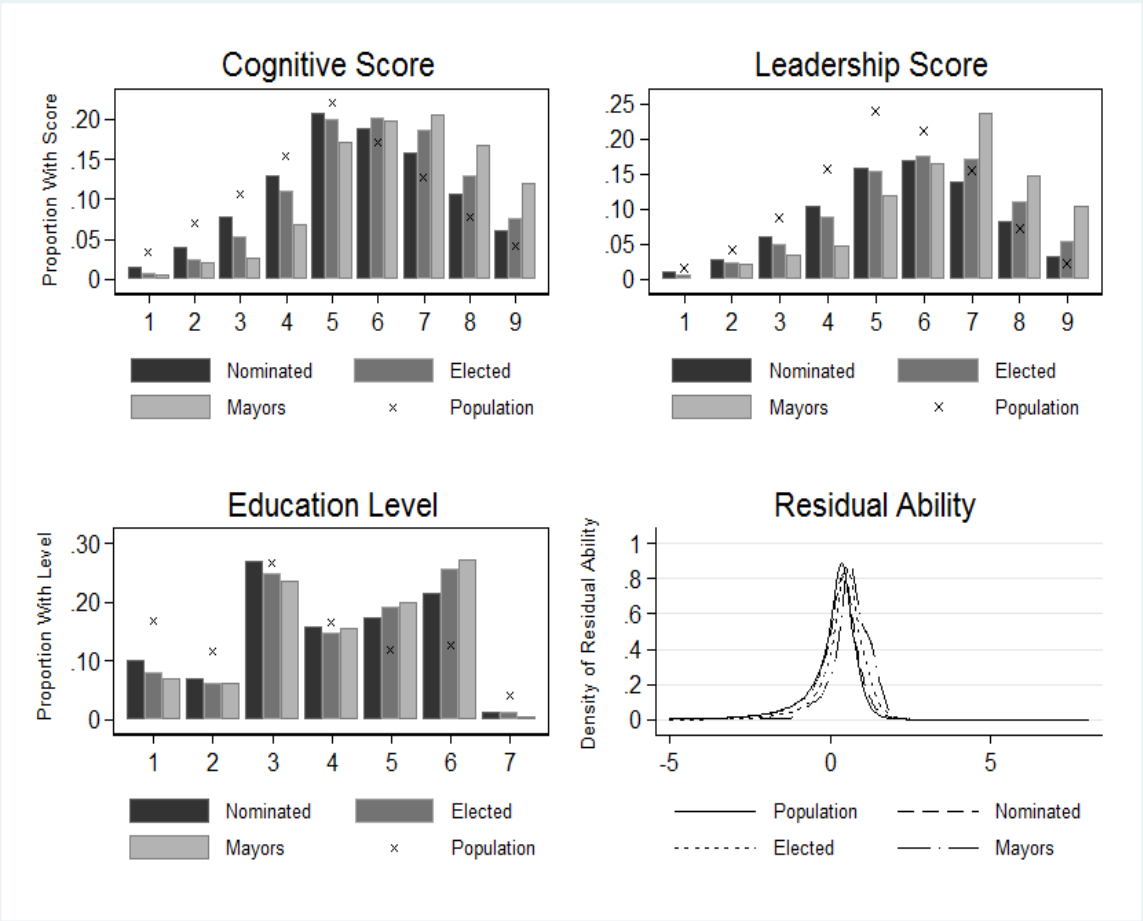

Figure 1: Distribution of ability measures in the population and among politicians

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, elected, and mayors. The two enlistment scores are shown in the top figures, the left bottom 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.

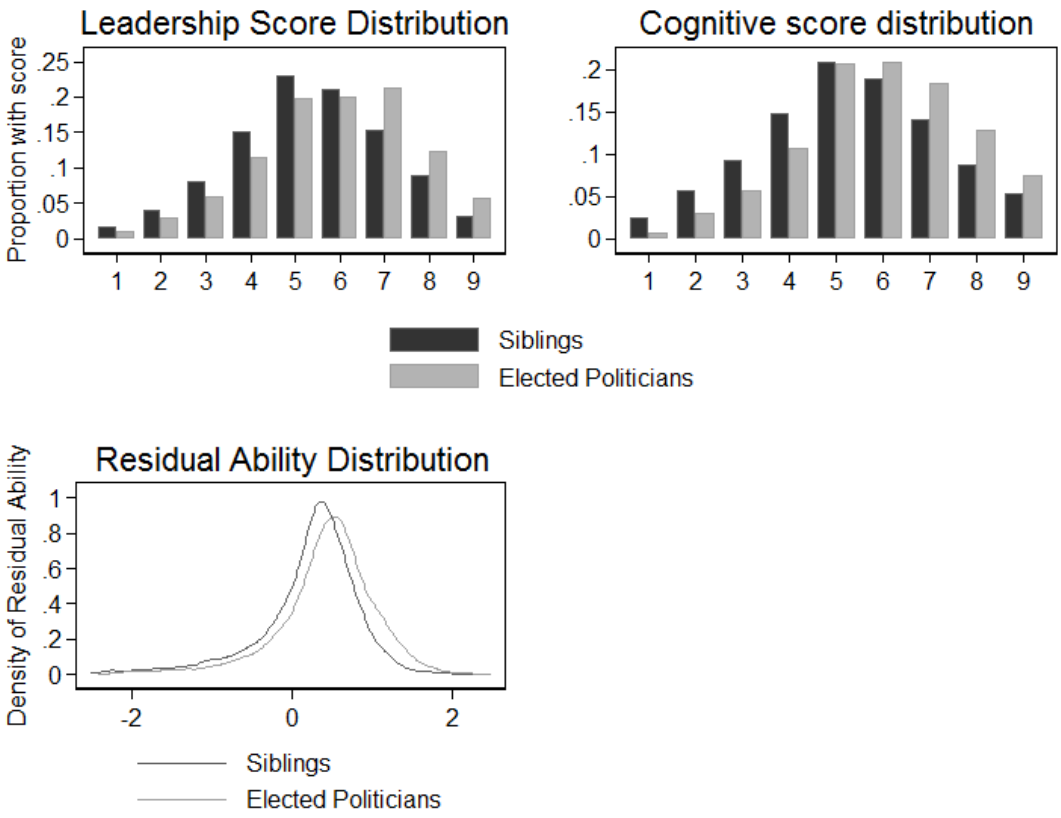

Figure 2: Distribution of ability measures among elected politicians and their siblings

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) and the cognitive score (right). The proportion of elected politicians with each score are represented by the light gray bars, while the dark gray bars show the proportions among the siblings. The bottom graph shows the distribution of the residual ability measure (z-score). The light gray line shows the distribution for the elected politicians and the darker gray line the distribution for the siblings. The data covers all politicians who were elected to a municipal council between 1982 and 2010, and who have at least one sibling. As all variables are constant over time, we use only one observation per individual. For the two enlistment scores, the sample is restricted to men in the 1951-1980 cohorts.

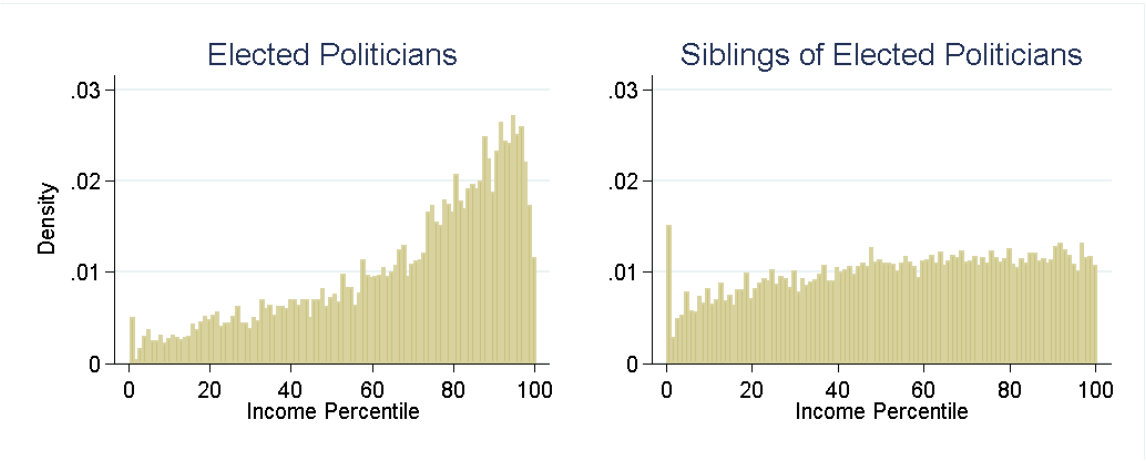

Figure 3: Distribution of elected politicians (left) and their siblings (right) across the percentiles of population income, 2011

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 working age (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.

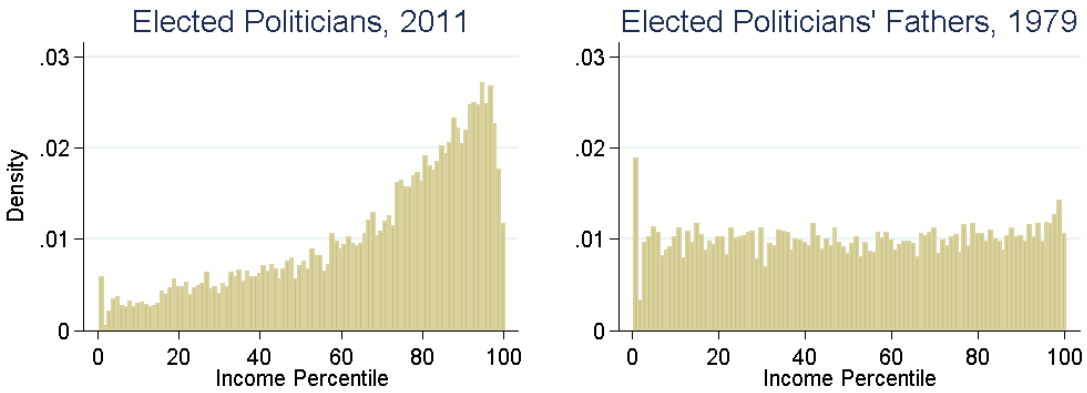

Figure 4: Distribution of elected politicians (left) and their fathers (right) across the percentiles of population income

Notes: The figure shows the distribution of elected politicians (left) and their fathers (right) across the percentiles of the Swedish income distribution. 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 working age in year 1979 (18-72), and politicians are excluded if we cannot find an earnings observation for their father in that year.

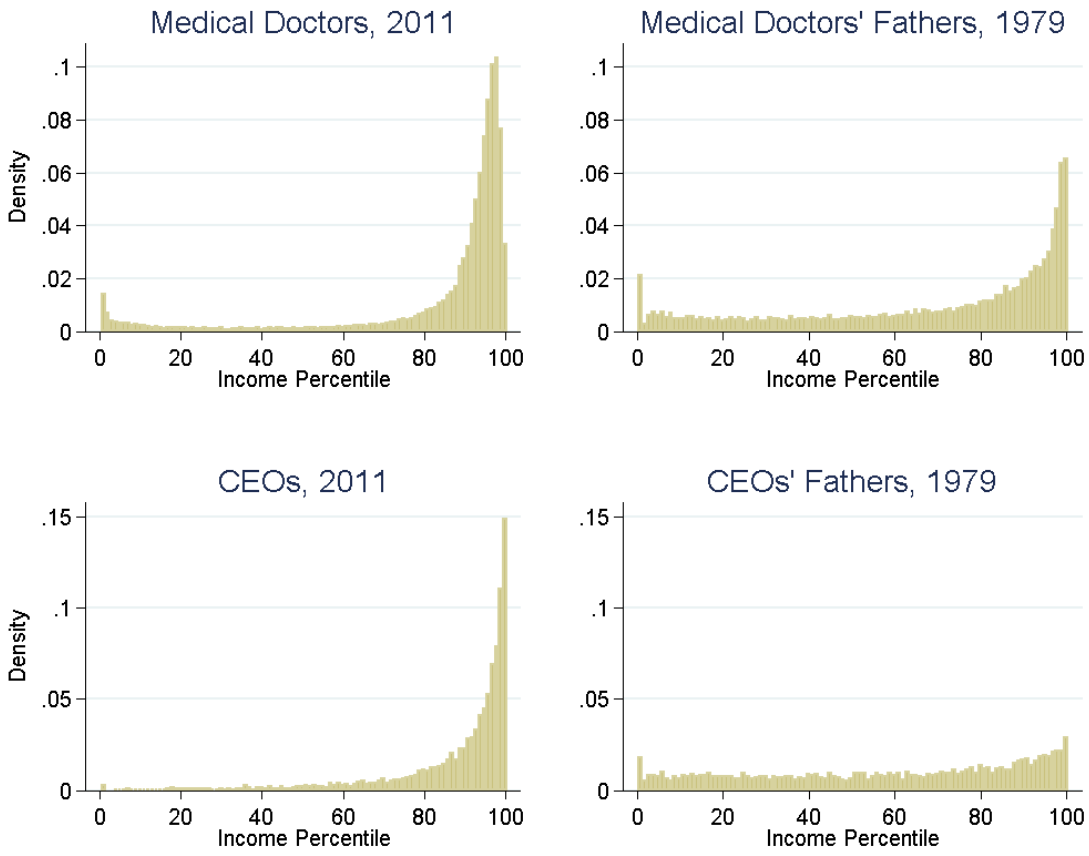

Figure 5: Medical doctors (upper left) and CEOs of firms with between 10 and 24 employees (lower left) and their fathers' (right) distribution across the percentiles of population income

Notes: The figure shows the distributions of the annual labor earnings of CEOs of firms with between 10 and 24 employees and medical doctors (left) and the corresponding distributions among their fathers (right). 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.

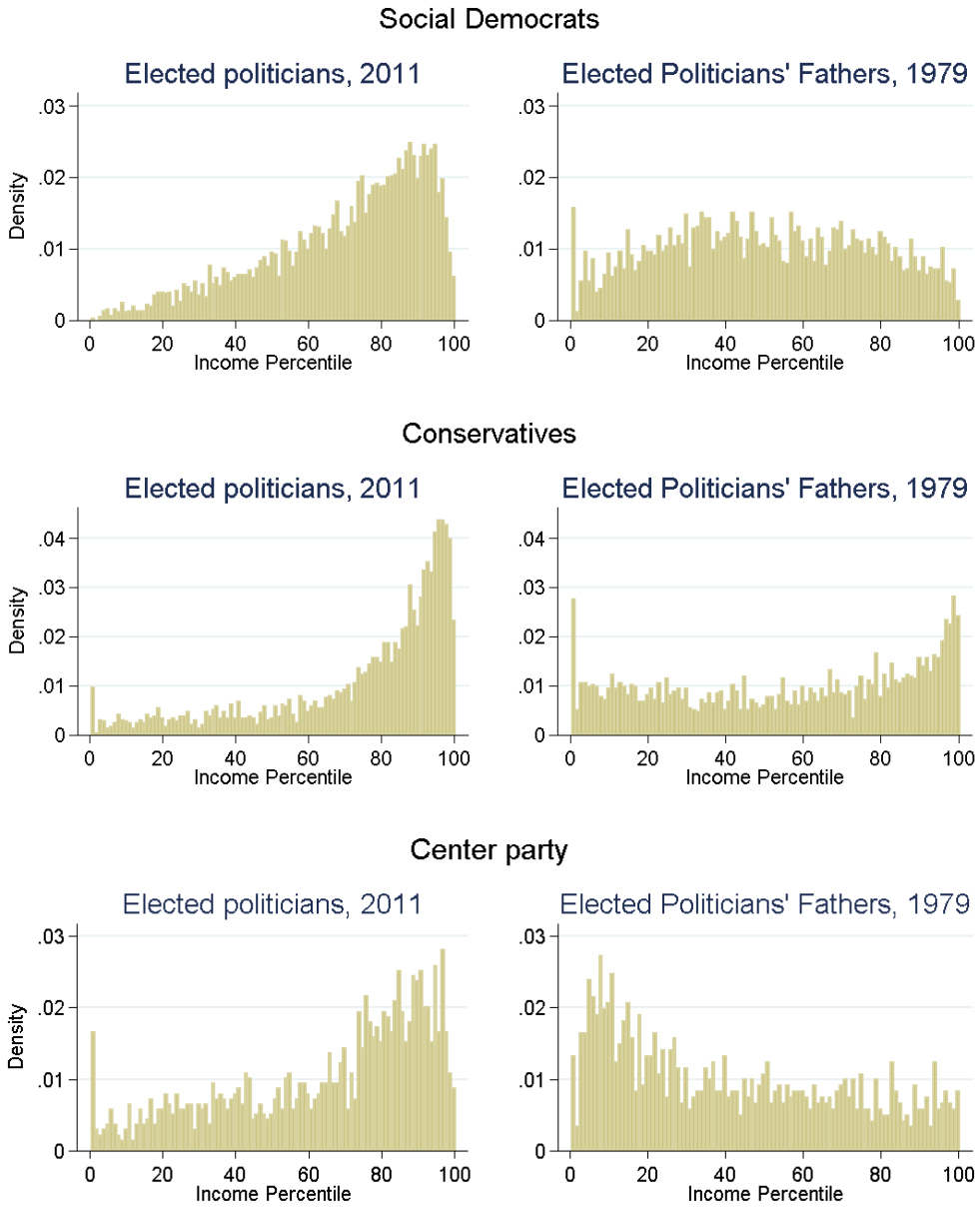

Figure 6: Distribution of elected politicians (left) and their fathers (right) across the percentiles of population income

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

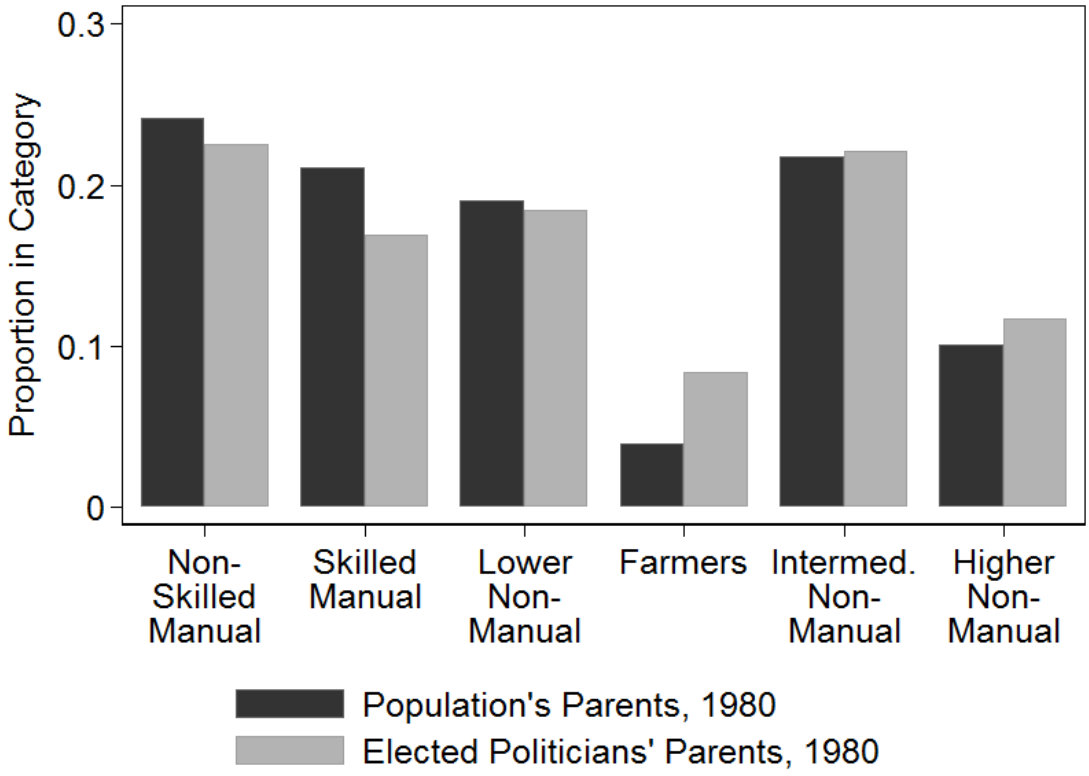

Figure 7: Distribution of 2010 politicians parents' social class, measured in 1980 and compared to the 2010 population's social class

Notes: The figure shows the distributions across six social classes of parents for elected politicians (light gray bars) and the population (dark gray bars). The height of the bars show 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 nominated politicians, elected politicians, and mayors

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 cognitive score and leadership 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 each category of politicians the average of the same competence measure in the working-age population in the same municipality and election year. In each graph, we display the selection index for three categories of politicians. The solid line shows the index for all politicians nominated on an electoral ballot in that municipality. The long dashed line shows the index for politicians elected to the municipal council, and the short dashed lines show the index for mayors. The data includes all election years from 1991-2010.

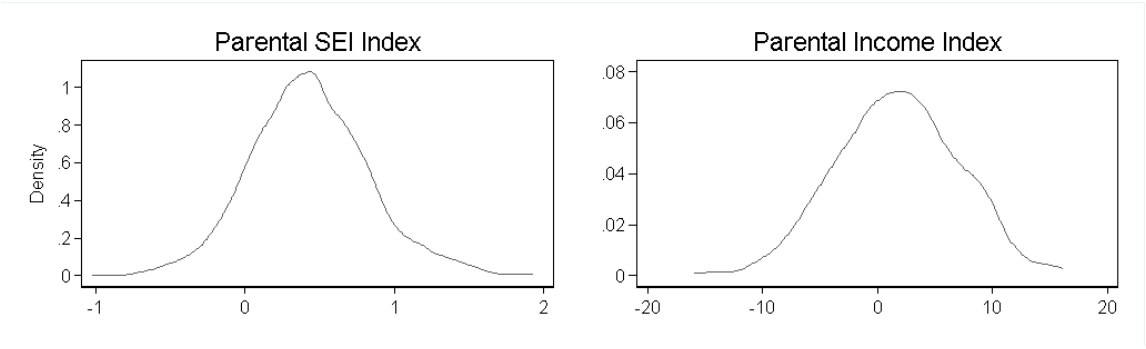

Figure 9: Distribution of representation indices across municipalities

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 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.

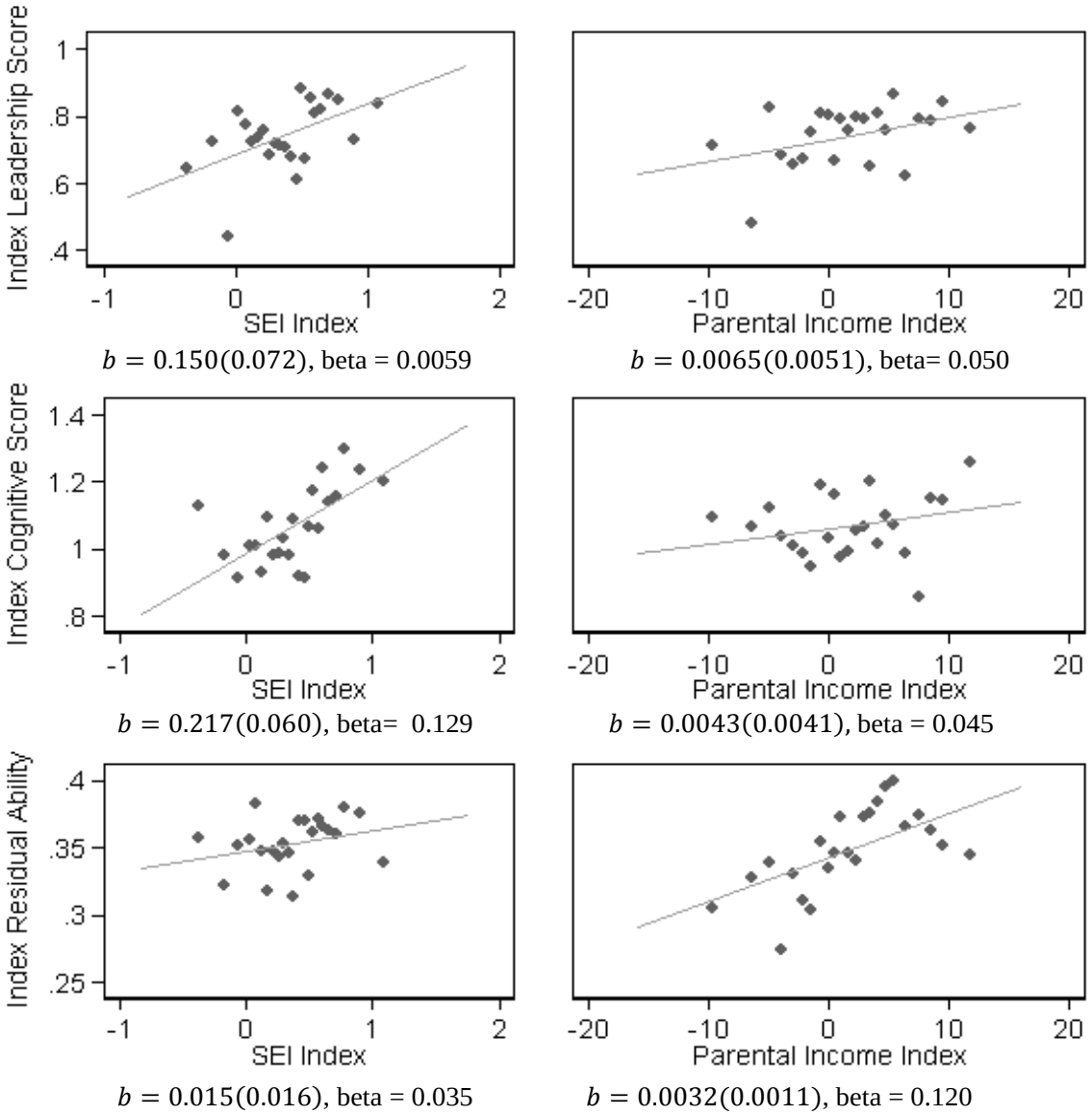

Figure 10: Correlations between municipal indices of representation and selection

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.

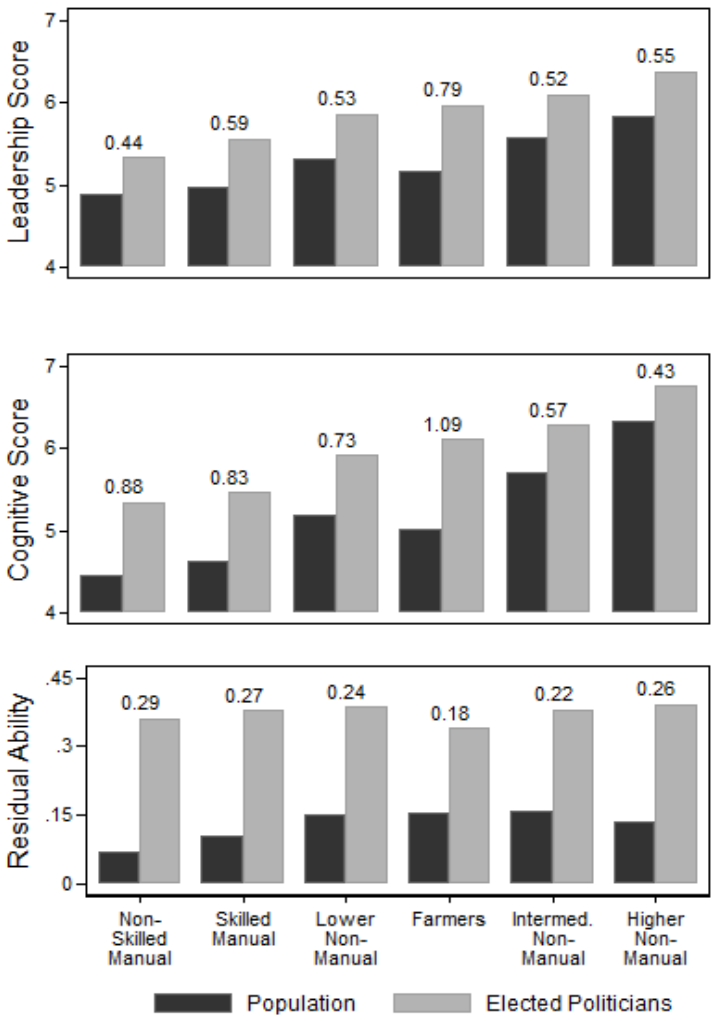

Figure 11: Ability by parental occupational class among politicians and in the general population

Notes: The figure shows the means of the three ability variables by social class of the politician's parents (x-axis). Means are shown for elected politicians (light gray bars) and the working-age population (dark gray bars). The number above each pair of bars shows the absolute difference between the two means in each parental social class. The top (middle/bottom) graph shows the means for leadership score (cognitive score/residual ability measure). For each elected politician, or individual in the population, we find their parents in 1980 and measure their social class (x-axis). Only the parent with the highest social class is included in the data. The data are pooled for all politicians elected in 2002, 2006, and 2010. The general population is sampled for these same years. For the cognitive and leadership scores, the sample is restricted to men in the 1951-1980 cohorts.

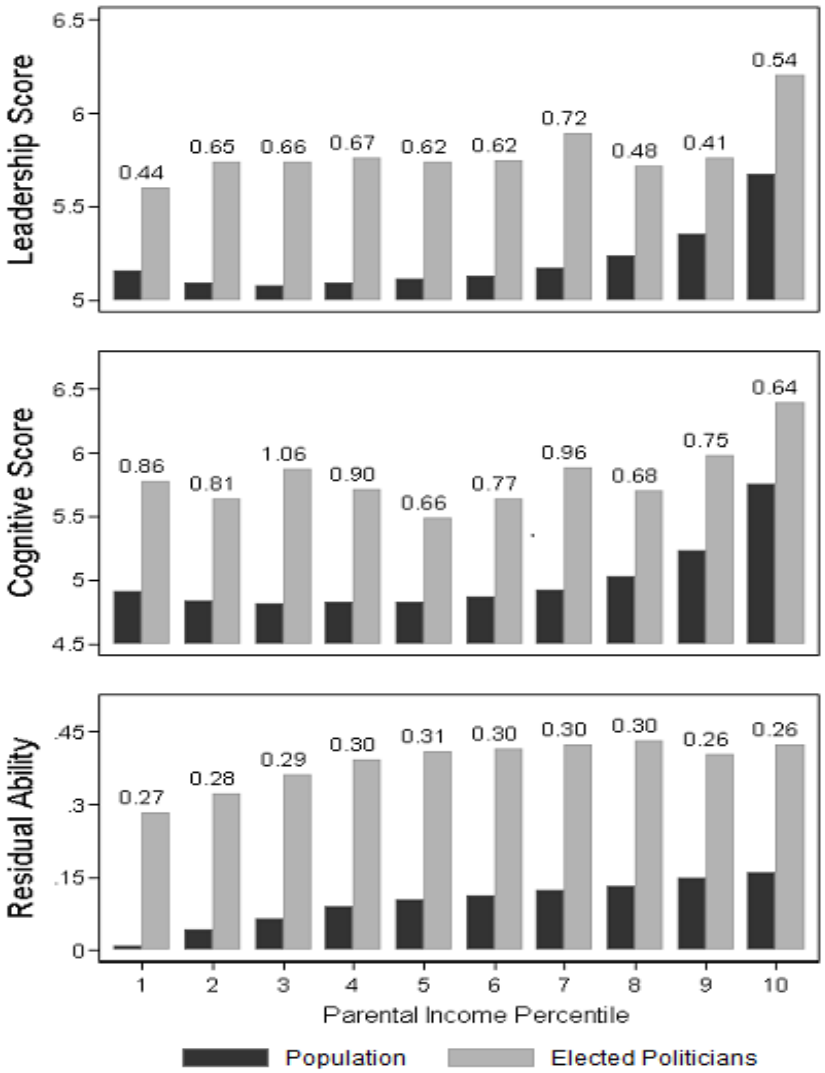

Figure 12: Ability by parental income decile among politicians and in the general population

Notes: The figure shows the means of the three ability variables for elected politicians (light gray bars), and for the working-age population (dark gray bars), by parental income decile. Data are drawn from the election years in 2002-2010. The general population is sampled for these same years. For the cognitive and leadership scores, the sample is further restricted to men the 1951-1980 cohorts. For each elected politician, or individual in the population, we find their parents in 1979 and measure the parents decile in the working-age population distribution of annual earnings (x-axis). Only the parent with the highest decile is included in the data.

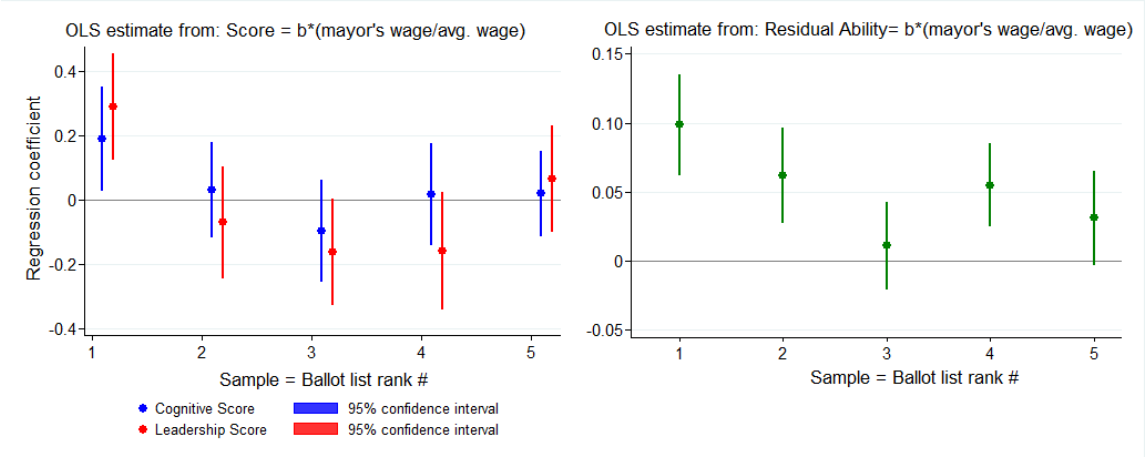

Figure 13: Estimated bivariate relationships between politician ability and the mayor's wage relative to the municipal population, by electoral ballot rank

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 run in sub-samples based on electoral ballot rank. These sub-samples are denoted on the x-axis. In the left graph, the cognitive score (blue color) and the leadership score (red 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.

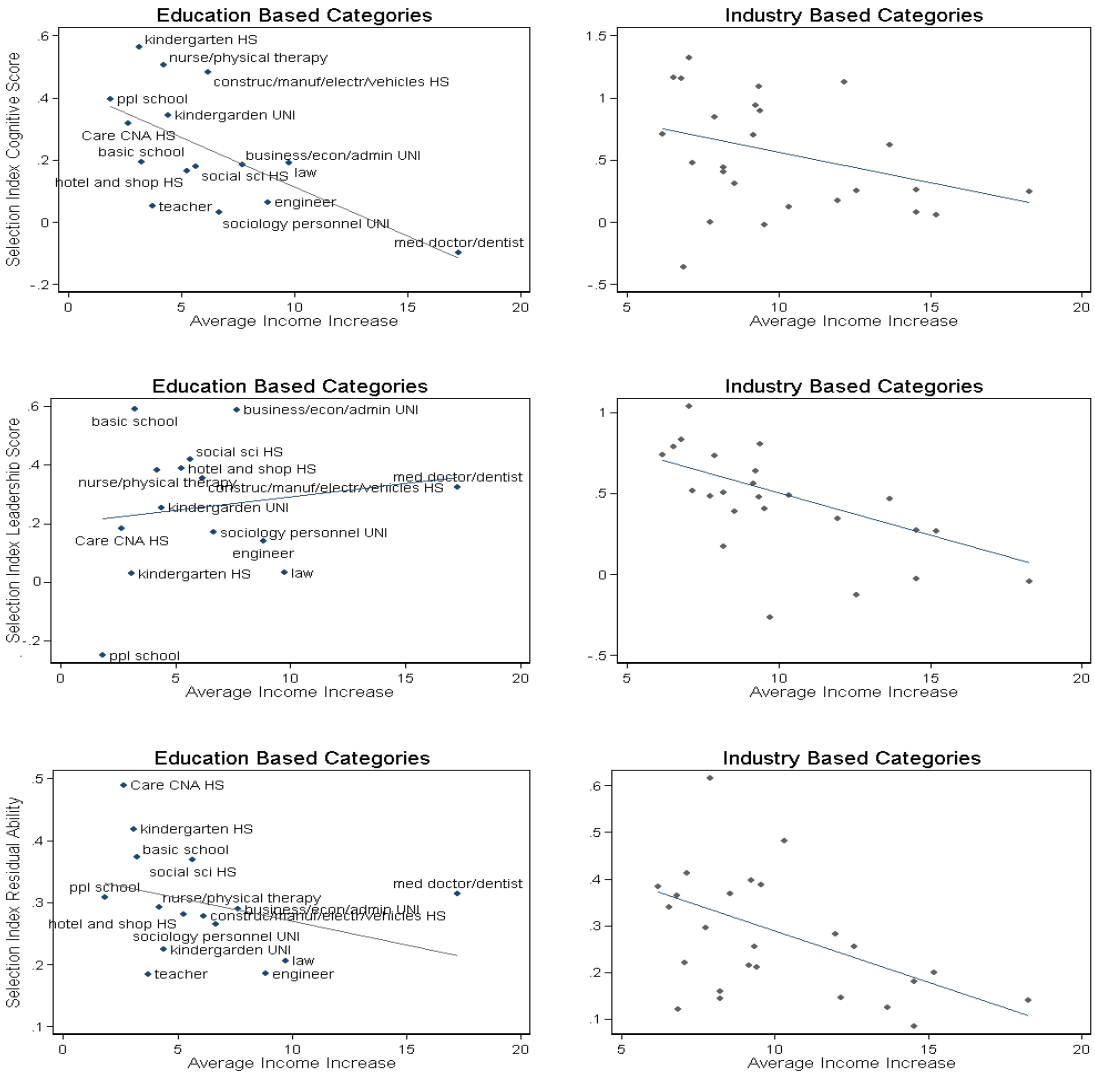

Figure 14: Ability indexes and earnings-tenure profiles

Notes: Elected municipal councilors during are divided into occupation, either by education (the left plots) or by 1-digit industry sector (the right plots). The x-axes of the left (right) column plots shows the average increase in annual earnings, in 1000 SEK, by education (sector), while the y-axes 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). People 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 regression. The cognitive and leadership scores are restricted to 1951-1980 cohort men.

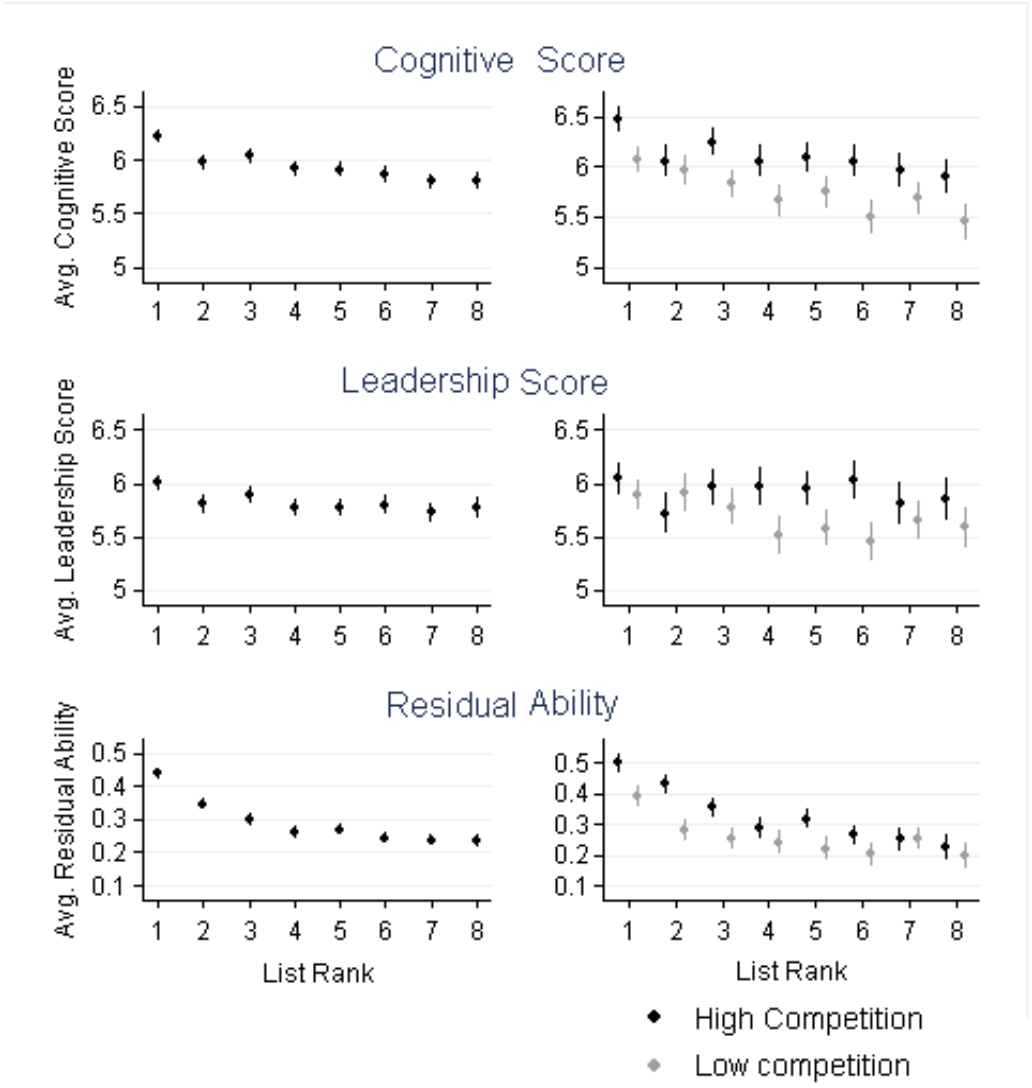

Figure 15: Average ability by party list rank

Notes: The figure shows the averages of three ability variables by ballot rank. The top (middle/bottom) graph shows the means for cognitive score (leadership 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 on 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

	Cognitive score	Leadership score	Years of education
Cognitive score	1		
Leadership score	0.338	1	
Years of education	0.511	0.3	1
Residual ability	0.138	0.171	0.065

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
Cognitive score (1-9)	5.1	1.9	5.5	5.9	6.3	6.4
Leadership score (1-9)	5.2	1.7	5.5	5.9	6.4	6.7
Years of education	12.4	3.0	13.4	13.7	13.7	14.8
Residual ability (z-score)	0.03	0.96	0.12	0.38	0.65	0.72
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 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

	Cognitive score	Leadership score	Residual ability	Years of schooling	Labor earnings	Obs
Mayors	6.2	6.4	0.62	13.9	679.4	247
Municipal councilors	5.9	5.8	0.31	13.8	379.0	8870
CEOs ($10-24$ employees)	5.8	6.1	0.61	13.6	675.6	6825
CEOs ($25-249$ employees)	6.2	6.4	0.78	14.2	1046.2	6885
CEOs (≥ 250 employees)	6.7	6.8	0.88	15.4	1926.0	1470
Medical Doctors	7.4	6.5	-0.09	17.1	640.0	29514
Lawyers and Judges	6.8	6.5	0.51	17.0	568.0	5308
Economists	7.0	5.9	0.16	20.4	530.0	248
Political Scientists	6.8	5.8	0.41	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 make up other "elite" occupations in Swedish society. Columns 1 and 2 reports the means for the two enlistment scores, the cognitive score and the leadership 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 the number of individuals classified into each elite occupation. The data is 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 by party-internal career prospects

	Governing bloc				Opposition bloc			
	Largest party		Minority party		Largest party		Minority party	
	Elected	Top	Elected	Top	Elected	Top	Elected	Top
Leadership Score	0.66	1.64	0.59	1.00	0.32	0.87	0.87	0.98
Cognitive Score	0.93	1.73	1.33	1.14	0.86	1.05	1.34	1.14
Residual Ability	0.43	0.65	0.21	0.39	0.37	0.60	0.20	0.40
Observations	316		413		324		178	

Notes: The table shows the means of three ability variables by two categorizations, the political status of the party (in four columns), and within each of these columns, the hierarchical level of the politicians, either elected (left column) or top-ranked on the electoral ballots (right column). 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 parties in each category of political status. Data is drawn from the election years in the 1990s and 2000s (1991, 1994, 1998, 2002, 2006, and 2010). The sample is restricted to municipalities where one bloc, left or center-right, has held a seat majority in the municipal council in five consecutive election periods. The largest party in the governing bloc refers to the party in that bloc holding the largest proportion of seats in the municipal assembly, and all other parties in that bloc are denoted as "minority parties in the governing bloc". The opposition bloc refers to the bloc that holds a seat minority. To further differentiate career opportunities in each bloc, we drop blocs where any of the smaller parties has a vote share within 75% of the vote share of the largest bloc member in any of the three most recent elections before t. We also exclude parties with less than three seats. For the enlistment measures of competence, only male politicians in the 1951-1980 cohorts are included.

A Appendix: Supplementary Figures and Tables

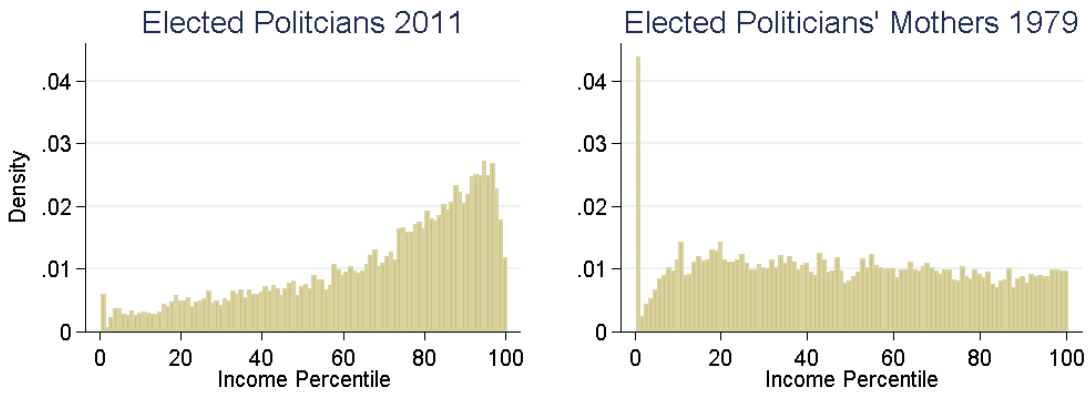

Figure A.1: Distribution of elected politicians and their mothers across the percentiles of population income

Notes: The figure shows the distributions of elected politicians (left) and their mothers (right) across the percentiles of the Swedish income distribution. 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 combine them into a "Leadership Score". The manual followed by the trained psychologist to make this assessment is tailored in a process that first scores the conscript on four sub-traits, listed in column 1 of the table, and then merge 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)).

Table A.2: Proportion of parties that have made appointments to positions of influence by our measure of (party-based) political career prospects

	Governing bloc		Opposition bloc	
	Largest party	Minority party	Largest party	Minority party
Mayor	97.50%	0.50%	0%	0%
Full time	94.50%	7.80%	16.80%	2%
Part time	44.00%	22.30%	39.50%	13.70%
Observations	200	242	205	95

Notes: The table shows the proportion of political parties that have made executive political appointments in four categories based on political status. Row 1 shows the proportions of parties that have appointed the mayor, row 2 the proportions that have appointed any full-time salaried politician, and row 3 shows the proportion of parties that have appointed any part-time salaried politician. We use data from 2006 and 2010, taken from a mandatory survey carried out by Statistics Sweden and collecting the personal ID codes for all indirectly appointed politicians. We further restrict the sample to municipalities where one bloc, left or center-right, has held a seat majority in the municipal council in five consecutive election periods. We also exclude parties that have less than three elected councilors, and we drop all parties in the blocs where any of the smaller bloc parties has a vote share within 75% of the largest bloc party in any of the three most recent elections before t. The four categories of political status are shown in the column titles. The largest party in the governing bloc is the party in that bloc that holds the largest proportion of seats in the municipal assembly, and all other parties in that bloc are considered "minority parties in the governing bloc". The opposition bloc is the bloc that holds the seat minority.

B Appendix: Formal Proofs

Proof of Proposition 1: The traits Y, P taking typical values (y, p) are independently and uniformly distributed in the space $T = \left[0, \bar{Y}\right] \times \left[0, \bar{P}\right]$, where the bounds $\bar{Y}, \bar{P}$ are assumed finite but arbitrarily large. The joint distribution is

$$\Pr\left(Y < y, P < p\right) = F_{YP}\left(y, p\right) = \frac{yp}{\bar{Y}\bar{P}} \ ,$$

with joint density $f_{YP}\left(y, p\right) = \frac{1}{\bar{Y}\bar{P}}$ and marginal densities $f_Y\left(y\right) = \frac{1}{\bar{Y}}$ and $f_P\left(p\right) = \frac{1}{\bar{P}}$.

The pool of volunteers is the set $A = \left\{(y, p) \in T | y < \frac{a+p}{\psi}\right\}$, where $a = \pi w$ and $\psi = \gamma\left[\pi\left(1 - \delta\right) + \delta\right]$. If $\bar{P} \leq \psi\overline{Y} - a$, the volunteer pool is a trapezoid with upper side $\bar{P}$, and right side $p = y - a$. If $\bar{P} > \overline{Y} - a$, the right side is $y = p + a$ for $p < \psi\overline{Y} - a$, and it is $\overline{Y}$ for $p \geq \psi\bar{Y} - a$.

To show that higher wages increase the average competence $E\left(Y | \left(y, p\right) \in A\right)$ in the volunteer pool, we need to show average competence increases in a. Note this will also imply average competence increases in π whenever $\delta \to 1$. In order to show that higher tenure returns γ decrease average competence, we need to show average competence decreases in ψ. Since a and ψ play opposite roles (given the self-selection condition $y < \frac{a+p}{\psi}$, we restrict attention to the comparative statics of a.

To do so, we divide $[0, \overline{P}]$ in two intervals $[0, P_1]$ and $[P_1, \overline{P}]$, where $P_1 = \overline{P}$ if $\overline{P} \leq \psi\overline{Y} - a$. Otherwise $P_1 = \psi\overline{Y} - a$.

Define:

$$h(a, p) = \begin{cases} \frac{a+p}{\psi} & if \quad \frac{a+p}{\psi} \leq \overline{Y}; (\text{here } p \in [0, P_1]\,) \\ \bar{Y} & if \quad \frac{a+p}{\psi} > \overline{Y}; \ (\text{here } p \in \left[P_1, \bar{P}\right]\,) \end{cases} \tag{5}$$

and

$$h_a(a, p) = \begin{cases} \frac{1}{\psi} & if \quad \frac{a+p}{\psi} \leq \overline{Y}; (\text{here } p \in [0, P_1]\,) \\ 0 & if \quad \frac{a+p}{\psi} > \overline{Y}; (\text{here } p \in \left[P_1, \bar{P}\right]\,) \ . \end{cases} \tag{6}$$

Therefore, the measure of A is $\int_{P=0}^{\bar{P}} \int_{Y=0}^{h(a,p)} f_{YP}\left(y, p\right) dydp$.

Average quality in the volunteer pool is

$$E[Y | \left(y, p\right) \in A] = \frac{\int_0^{\overline{P}} \int_0^{h(a,p)} y \, dydp}{\int_0^{\overline{P}} \int_0^{h(a,p)} dydp} = \frac{1}{2} \frac{\int_0^{\overline{P}} h^2(a, p) \, dp}{\int_0^{\overline{P}} h(a, p) \, dp} \ .$$

We now take the derivative of the term $\frac{\int_0^{\overline{P}} h^2(a,p) \, dp}{\int_0^{\overline{P}} h(a,p) \, dp}$:

$$2\frac{dE[Y|\,(y,p)\in A]}{da} = \frac{2[\int_0^{\overline{P}} h(a,p)h_a(a,p)\,dp][\int_0^{\overline{P}} h(a,p)\,dp] - [\int_0^{\overline{P}} h^2(a,p)\,dp][\int_0^{\overline{P}} h_a(a,p)\,dp]}{\left(\int_0^{\overline{P}} h(a,p)\,dp\right)^2}$$

$$= \frac{2\left(\int_0^{P_1} h(a,p)\,dp\right)\left(\int_0^{P_1} h(a,p)\,dp + \int_{P_1}^{\overline{P}} h(a,p)\,dp\right)}{\left(\int_0^{\overline{P}} h(a,p)\,dp\right)^2}$$

$$- \frac{\left(\int_0^{P_1} h^2(a,p)\,dp + \int_{P_1}^{\overline{P}} h^2(a,p)\,dp\right)\left(\int_0^{P_1}\,dp\right)}{\left(\int_0^{\overline{P}} h(a,p)\,dp\right)^2},$$

where we have used the fact that $h_a(a,p) = 0$ for $p \in \left[P_1, \bar{P}\right]$. After substantial algebra, the term on the RHS of the last equality, can be expressed as

$$= \frac{P_1^2}{\psi^3}\frac{\left(\frac{1}{2}\frac{1}{\psi} - \frac{1}{3}\right)P_1^2 + 3a\,(a+P_1)}{\left(\int_0^{\overline{P}} h(a,p)\,dp\right)^2}$$

$$+(\overline{P}-P_1)(a+P_1)P_1\frac{1}{\psi^2}\frac{\frac{1}{\psi}\,(2a+P_1) - (a+P_1)}{\left(\int_0^{\overline{P}} h(a,p)\,dp\right)^2},$$

where the first term is strictly positive and the second greater or equal to zero, yielding $\frac{dE[Y|Y\leq\frac{a+P}{\psi}]}{da} > 0$.■

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

Negative screening. Here, the party selects type $(0,0)$ 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 $E\,(y) = \frac{\bar{Y}}{2}$, we analyze two cases.

Case (i) $\bar{P} \leq \psi\overline{Y} - a$. This means $\frac{a+p}{\psi} \leq \bar{Y}$ for all $p \leq \bar{P}$. Then, the volunteer set is bounded at the right by the entry line $p\,(y) = (\pi(1-\delta)+\delta)\,\gamma y - \pi w$, or

$y = \frac{p+\pi w}{(\pi(1-\delta)+\delta)\gamma} = \frac{a+p}{\psi}$.

$$\begin{aligned} E(y_A) &= \frac{\int_0^{\overline{P}} \int_0^{\frac{a+p}{\psi}} y\, dydp}{\int_0^{\overline{P}} \int_0^{\frac{a+p}{\psi}} dydp} = \frac{\int_0^{\overline{P}} \frac{1}{2}\left(\frac{a+p}{\psi}\right)^2 dp}{\int_0^{\overline{P}} \frac{a+p}{\psi} dp} = \frac{\int_0^{\overline{P}} \frac{1}{2}\frac{a+p}{\psi} dp}{\int_0^{\overline{P}} dp} \\ &= \frac{1}{2}\frac{a+p}{\psi} < \frac{\bar{Y}}{2} = E(y) \ , \end{aligned}$$

by the assumption that $\frac{a+p}{\psi} \leq \bar{Y}$.

Case (ii) $\bar{P} > \psi\overline{Y} - a$. This means $\frac{a+p}{\psi} \leq \bar{Y}$ for all $p \leq P_1 = \psi\bar{Y} - a$, but $\frac{a+p}{\psi} > \bar{Y}$ for $p > P_1$. Then we have

$$\begin{aligned} E(y_A) &= \frac{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} y\, dydp + \int_{P_1}^{\overline{P}} \int_0^{\bar{Y}} y\, dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp + \int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp} = \\ &= \frac{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} y\, dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp} \frac{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp + \int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp} \\ &\quad + \frac{\int_{P_1}^{\overline{P}} \int_0^{\bar{Y}} y\, dydp}{\int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp} \frac{\int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp + \int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp} \\ &= \frac{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} y\, dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp}\alpha + \frac{\int_{P_1}^{\overline{P}} \int_0^{\bar{Y}} y\, dydp}{\int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp}(1-\alpha) \ , \end{aligned}$$

where $\alpha \equiv \frac{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp}{\int_0^{P_1} \int_0^{\frac{a+p}{\psi}} dydp + \int_{P_1}^{\bar{P}} \int_0^{\bar{Y}} dydp}$. Thus,

$$\begin{aligned} E(y_A) &= \frac{\int_0^{P_1} \frac{1}{2}\left(\frac{a+p}{\psi}\right)^2 dp}{\int_0^{P_1} \frac{a+p}{\psi} dp}\alpha + \frac{\int_{P_1}^{\overline{P}} \frac{1}{2}\bar{Y}^2 dp}{\int_{P_1}^{\bar{P}} \bar{Y} dp}(1-\alpha) \\ &= \frac{1}{2}\left(\frac{a+p}{\psi}\alpha + \bar{Y}(1-\alpha)\right) < \frac{\bar{Y}}{2} \end{aligned}$$

since $\frac{a+p}{\psi}$ is only defined over values of p such that $\frac{a+p}{\psi} < \bar{Y}$.

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 $\left(\bar{P}, \min\left\{\frac{\bar{P}+w\pi}{(\pi(1-\delta)+\delta)\gamma}, \bar{Y}\right\}\right)$.■