Quota vs Quality?
Long Run Impacts of a Gender Quota

Ursina Schaede & Ville Mankki

October 29, 2021

Click here for the most recent version

Abstract

We investigate the potential equity-efficiency trade-offs from quotas by studying the effect of abolishing a 40% quota for male primary school teachers on pupils' long run outcomes in Finland. The quota had advantaged academically lower scoring male applicants in university admissions, and its removal cut the share of men among new teachers by half. We show that combining this reform with the timing of union-mandated teacher retirements creates quasi-random variation in the local share of male teachers. We then document how quota-induced changes in local teacher gender composition affect pupils' outcomes in the long-run. Using comprehensive register data, we find that pupils exposed to a higher share of male quota teachers during primary school have higher educational attainment and labor force attachment at age 25. Pupils of both genders benefit similarly from exposure to male quota teachers. Our findings are consistent with male quota teachers contributing valuable qualities to school environments that are not sufficiently captured by the teacher selection process in absence of the quota.

We are grateful to David Yanagizawa-Drott, Joachim Voth, Claudia Goldin and Ulf Zölitz for their guidance throughout this project. We thank Anne Sofie Beck Knudsen, Augustin Bergeron, Anne Brenøe, Lorenzo Casaburi, Ana Costa-Ramón, Alessandro Ferrari, David Hémous, Andrea Hofer, Kristiina Huttunen, Mika Kortelainen, Ross Mattheis, Morten Olsen, Heather Sarsons, and Hanna Virtanen for helpful comments and suggestions. We thank Topias Jalo and his colleagues at Statistics Finland, as well as the administrative staff at the Universities of Tampere and Zurich for their support. We gratefully acknowledge financial support from the Yrjö Jahnsson foundation.

1 Introduction

Are affirmative action policies, such a as quotas, costly for economic efficiency? On one hand, when quotas in educational institutions and workplaces are binding, they require these organizations to admit applicants who would have been rejected otherwise. As quotas force a decision maker to deviate from her preferred ranking of candidates, the overall quality of successful applicants may be lower (Welch, 1976; Lundberg and Startz, 1983; Arcidiacono and Lovenheim, 2016). Cast in this light, affirmative action policies may achieve a distributional goal only at a cost of lower productivity.

In contrast to this reasoning, affirmative action policies may raise economic efficiency when evening out differences that arise from unequal opportunities that are unrelated to inherent potential ability: In this case, affirmative action measures may improve the allocation of talent in the economy (Hsieh et al., 2019; Coate and Loury, 1993; Becker, 1957).

Even in the absence of explicit discriminatory barriers, differential treatment of underrepresented groups may be desirable when selection processes operate under imperfect knowledge of candidates' potential abilities. This is the case in settings in which multiple dimensions of skill determine output, but selection criteria are limited to specific subsets of cognitive scores (Heckman et al., 2006; Deming, 2017; Bell et al., 2019; Siniscalchi and Veronesi, 2021). A similar reasoning applies to environments in which there are complementarities between groups that materialize in production, such that diversity itself augments productivity. Taken together, these countervailing forces indicate that whether a representation quota is costly or output-enhancing is ultimately an empirical question whose answer depends on the specifics of the setting.

In this paper, we document that a quota - beyond merely achieving a representational goal - raised efficiency in the relevant industry by correcting a mis-allocation of talent. Leveraging a unique policy experiment in Finland, we trace out how a quota-induced shift in the gender composition of an entire occupation affects output in the long run. We study how the termination of a quota for male primary school teachers in Finland influenced pupils' educational outcomes and subsequent labor force attainment. Specifically, this quota had reserved 40% of slots in the final admissions round for primary school teachers studies at Finnish universities for male applicants, advantaging academically lower-scoring male candidates. The removal of the quota resulted in a sharp drop in the share of men among admits to primary school teacher studies from about 40% to 20% (Uusiautti and Määttä, 2013; Räihä, 2010).

Our identification strategy isolates exogenous variation in pupils' exposure to male quota teachers with a differences-in-differences instrumental variables (DiD-IV) framework: To instrument for the local teacher gender composition that pupils experience in primary school, we use the lifting of the quota together with local shocks to the demand for new teachers. These demand shocks arise from local teachers reaching the union-mandated teacher retirement age when they turn 60. Intuitively, municipalities in which teachers turn 60 while the quota is still in place will hire new teachers from a rookie teacher market with quota men, compared to municipalities whose teachers turn 60 just after the quota was abolished.[1] The exclusion restriction in this set-up requires that teacher retirements do not differentially affect pupil outcomes in the post-quota period except via teacher gender composition. Our empirical strategy addresses this by comparing pupils who experience similar exposure to new teachers via retirements, but face a different gender composition of those rookie teachers due to the lifting of the quota.

We start by outlining a simple conceptual framework of university admissions that derives conditions under which a representation quota results in lower or higher total ability of admitted candidates. When the selection criterion fully reflects candidate ability, introducing a quota comes at the cost of admitting less qualified candidates. In contrast, there may be efficiency gains from a quota when the selection criterion is negatively correlated with minority group status and thus insufficiently captures the potential ability of candidates: This can be the case both when the mapping between the selection criterion and ability differs across groups, and when there are complementarities in production between groups.

We then turn to examine the efficiency effects of the quota in the Finnish context empirically. First, we document how the lifting of the quota affects the local gender composition of teachers at the municipal level: Once the rookie teacher cohorts that studied without the quota enter the market for rookie teachers in 1994, each retiring teacher is 20 percentage points less likely to be replaced with a male teacher relative to the quota period. These changes in the local teacher gender composition are accompanied by small, but noisily measured, increases in local teachers' average academic scores, consistent with the notion that lower scoring men are no longer admitted to primary teachers studies once the quota is abolished.

We proceed to study how these changes affect pupils, using comprehensive register data from 1988 to 2018 to trace out pupils' education and labor market pathways until age 25. Using the timing of teacher age-based retirements as an instrument for the local teacher gender composition, we show that pupils exposed to a higher share of male teachers via the quota results in higher educational and labor force attainment in the long run: At age 25, these children are .09 SD more likely to be a student or actively participating in the labor market for a 1 SD increase in the share of male quota teachers.[2] In terms of educational attainment, we observe a shift to higher qualifications throughout the educational attainment distribution: For practically-oriented vocational degrees, pupils are more likely to have additional advanced qualifications instead of a basic three year degree. For academic tracks, pupils are more likely to have obtained a university level BA degree by age 25.

To better understand from where these differences in long term outcomes originate, we use rich register data to closely track pupils' educational trajectory after finishing compulsory schooling. We use records from the nationally-organized allocation of education slots after compulsory schooling, for which pupils can put in up to five preferred institution choices. We find that male quota teachers have important effects at this pivotal moment: Pupils exposed to a higher share

[1] We use the term 'quota men' to refer to those men that were only able to become a teacher because of the quota. Throughout the paper, we refer to teachers turning 60 as 'retirement'.

[2] 1 SD in the share of male quota teachers corresponds to 6.5 percentage points.

of male quota teachers are more likely to directly apply for and enroll in continued education. As pupils apply more in line with their academic abilities, they are more likely to obtain one of their top two choices, and less likely to end up without a spot in continued education. These patterns translate into higher enrollment rates in post-compulsory education at age 16.

A central question that arises from our results is through which characteristics quota men – relative to marginal female teachers pushed out by the quota – make a difference. While data limitations preclude us from linking teachers and pupils to classrooms, we can still shed light on the different channels through which male quota teachers have an impact.

First, via a same-gender role model channel, male quota teachers could raise specifically boys' aspirations and achievement, thus helping to close academic achievement gaps between male and female pupils. We can rule out such a role model channel that would purely operate through male teachers setting an example for boys via same-gender identity: We find no evidence for boys' educational attainment being more affected from exposure to male quota teachers relative to girls, and none of our main effects differ systematically or significantly by pupil gender. In a similar vein, we do not find evidence for male teachers substituting for male role models at home, as effects are not more concentrated among pupils that grow up in a single parent household.

Second, quota men could inspire pupils to pursue different educational fields. Our estimates indicate that pupils of either gender move away from more gender-neutral education choices and become more likely to choose a STEM field instead. We also examine whether exposure to more male quota teacher inspires boys to take up education related fields themselves. However, we find that pupils of either gender are not more likely to have chosen education-related or teaching fields by age 25.

Third, complementarities between male and female teachers (i.e. by specializing according to comparative advantage) could result in better outcomes for all pupils. The Finnish primary school system is characterized by extensive collaboration between teacher colleagues, both in school-wide curricula design and preparation of classes, as well as in actual teaching. Our analysis of the role of complementarities between male and female teachers estimates separate effects depending on the gender composition of the local teacher team. We find that the benefits of adding an additional male teacher are similar in magnitude between places with few male teachers and places where the share of men among colleagues is already high. However, a lack of precision prevents us from reaching decisive conclusions about whether our results are driven by quota men having inherently higher teaching ability than marginal women, or whether complementarities make teacher teams more productive.

We posit that the evidence in our setting is consistent with male quota teachers contributing positive qualities to the school environment that are not sufficiently captured by the selection criteria in absence of the quota. We document that the first order effects of the misalignment between these criteria and male quota teachers' effectiveness on the job dwarf any potential same-gender role model effects. While the quota thus neither closed educational attainment gaps between male and female pupils, nor mitigated the gender-segregation of teaching fields in the long term, it succeeded in selecting male teachers with valuable characteristics, resulting in better outcomes for pupils of either gender.

This study makes two main contributions. First, by examining the efficiency effects of a quota, we relate to recent work on affirmative action in university admissions documenting that marginal admits benefit similarly from access to selective colleges in the United States relative to marginal candidates pushed out (Bleemer, 2021; Bleemer, 2020; Black et al., 2020). Empirical evidence on how quotas specifically impact productivity-related measures has almost exclusively focused on quotas for women in board rooms, documenting negative or neutral effects on firm performance in the short run (Ahern and Dittmar, 2012; Matsa and Miller, 2013; Eckbo et al., 2021; Ferrari et al., 2021). We contribute to these papers by cleanly documenting that a quota – applied at university entry – improves output in the relevant sector in the long run. Our results also highlight a broader point, further illustrated in our conceptual framework: Selection of candidates based on academic scores and interviews, among the most widely used methods to assess applicants, can miss out on important dimensions of talent. A quota can mitigate such misalignment in settings when prior academic achievement and admissions committees' evaluations are imperfect predictors of effectiveness on the job.

Second, we shed light on the channels through which male quota teachers have a positive impact on pupils. Quasi-experimental studies studying the role of teacher gender have sought to explain effects via same-identity role model channels, such that a same-identity teacher affects academic performance (Dee, 2007; Antecol et al., 2015; Lim and Meer, 2017; Carrell et al., 2010). Work on role model interventions has documented positive effects for female minority students (Porter and Serra, 2019; Breda et al., 2018; Kofoed et al., 2019). A key advantage of our study is that we can examine whether shifting the gender composition of an entire occupation via a quota has impacts on pupils via same-identity role model channels. We can thus address pressing policy questions regarding a quota's potential to change the gender-segregation of occupation categories for future generations, with pupils in our setting being directly exposed to an occupation that is diversified via a quota.

Shedding light on how trade-offs related to equitable representation targets materialize by assessing their effects on productivity is important – especially so at a time at which affirmative action policies are under scrutiny. While many countries around the world are aiming to further diversify occupation categories (UN, 2019; Long, 2019), there is also significant push back in other settings. Universities in the United States and elsewhere are facing increasing judicial challenges for admission policies that embrace broader measures of academic preparedness (Leman, 2021; Dhume, 2019). Our results suggest that carefully considering the correlates of ability and minority group status in selection processes provides a promising avenue to increase both equitable representation and economic efficiency.

The paper is structured as follows: The next section details the Finnish education and teacher training system, and examines the quota reform as well as related literature in more depth. We outline a brief conceptual framework in section 3 to show more formally under which conditions introducing a quota would be more costly or lead to positive effects on output. Section 4 explains our data sources and sample, followed by the empirical design in section 5. In section 6, we first examine the effects of the quota on teacher gender composition at the municipal level, before turning to a pupil-based panel. We document mechanisms in section 7 and robustness checks in section 8. The final section concludes.

2 Context

2.1 Primary school teachers in Finland

Finland has been among the top scoring countries for multiple rounds of PISA evaluations, leading to considerable international attention paired with efforts to copy best practices from the Finnish education and teacher training system (Malinen et al., 2012; Niemi et al., 2016). Finland prides itself in having an equitable school system that aims at equalizing opportunities. Throughout comprehensive school, which spans grades 1-6 in primary school and grades 7-9 in middle school, there is no academic tracking, no standardized testing, very rare grade retention, and inclusive learning (Sahlberg, 2021).

Due to being one of the most competitive degrees in university admissions, primary school teachers enjoy high social status (Finnish National Agency for Education, 2018). While salaries are on par with OECD average, active teaching hours are comparatively low (Sahlberg, 2021). Primary school teachers are municipal employees that are hired by local schools, and are part of a powerful teacher union that sets both salary schedules and retirement at age 60 (Kivinen and Rinne, 1994). A national curriculum outlines broad learning goals and teachers within and across schools collaborate in designing detailed learning plans under the supervision of municipal education authorities (Sahlberg, 2021; Sahlberg et al., 2019).

In contrast to the United States, primary school teachers can spend several years with one cohort of pupils as their main classroom teacher covering all subjects. However, primary school teachers are also actively embedded in their work environment through extensive collaboration with their colleagues, both in curriculum design, preparing lessons and school wide activities, as well as in active teaching (Sahlberg, 2021). Co-teaching, i.e. the joint teaching of two classrooms by two teachers together, is frequently practiced with more than half of primary school teachers co-teaching at least once a week (Saloviita, 2018).

Pupils in our setting are thus experiencing and interacting regularly with the teacher body of their entire school. Conducting the analysis at the municipal level takes into account any spill-overs that arise from teachers collaborating within and across schools.

2.2 Primary school teacher training and the quota reform

The first teacher training institutes in Finland were founded in the mid-1800s, and offered training separate by gender. In 1881, new education decrees allowed for co-education for children attending municipal primary schools as long as sufficient instruction in handicrafts could be guaranteed, de facto leading to "differentiation between male and female elementary school teachers and a quota system in teacher training" (Sysiharju, 1987). Universal, co-educational, and compulsory primary school spanning grades 1-6 was introduced shortly after national independence, in 1921 (Niemi et al., 2016). Due to the number of study slots by gender allocated in teacher training institutes, Finland produced roughly 4 male for every 6 female primary school teachers (Sysiharju, 1987; Uusiautti, Määttä, et al., 2013; Räihä, 2010). While the number of primary school teachers ballooned from 6,800 to 25,000 from the 1920s to the late 1960s, the share of male teachers remained stable at 41% (Sysiharju, 1987).

In the context of major educational reforms in the 1970s, the original teacher training institutes were dissolved and primary teacher education was exclusively taught as a research-oriented, five year masters' degree at universities (Niemi et al., 2016). With an acceptance rate of about 10%, primary school teaching has been and still is among the most competitive degrees in the country and applicants often apply multiple years in a row until they are successfully admitted (Tirri, 2014; Uusiautti and Määttä, 2013).

Admissions throughout our study period closely followed the main principles established in those reforms, including that "the Ministry of Education maintained the sex quota system for the training of classroom teachers" (Sysiharju, 1987): In a first step, applicants were ranked according to a score that mainly considered candidates' grades in the matriculation exam, a nationally graded high school exit exam that qualifies for university admissions, with a few additional points given for candidates' extra-curricular activities. The highest ranking candidates were invited to an in-person second round (Uusiautti, Määttä, et al., 2013). In this second step of the selection process, candidates' teaching qualities were evaluated independently by a faculty board through an extensive host of exercises, including (among others) a live teaching exercise in front of real pupils, a pedagogical analysis of this practice lesson, and in-depth interviews (Räihä, 2010; Uusiautti, Määttä, et al., 2013). The highest ranked candidates in the second step were admitted to study primary teacher education according to the number of available study slots. Within this process, the quota suggested that universities should reserve 40% of study slots to male applicants.(Sysiharju, 1987; Räihä, 2010).

Despite the quota being a relatively soft measure since it did not force universities to admit male quota candidates, about 40% of active primary school teachers in the country were male until the mid-1980s (Sysiharju, 1987). Documentation on candidates invited to the second step of selection show that about 40% were men until 1988, and their share dropped to 20% in 1989. Statistics on final admissions in 1983 indicate that men constituted about a quarter of total applicants, and had a success rate of 15.5%, while that of women was around 8.5% (Sysiharju, 1987).[3]

The quota was abolished for the cohort applying to university in the fall of 1989 (thus graduating from primary school teacher studies in 1994). The Equality Ombudsman (Tasa-

[3]Uusiautti and Määttä (2013) cite a statistic from the University of Lapland in 1978, where about 3 applicants were invited per available study slot. In light of the quota expiring, a working group of the Ministry of Education recommended universities to invite at least four candidates per available study slot for the second round to ensure enough diversity among selected teachers without the quota (Etelä Suomen Sanomat, 1988). Mankki et al. (2020) examine scores received in the second step of the selection process in more recent years (after the quota), and show that male candidates are much more likely than women to receive the highest score.

Arvovaltuutettu), who was in charge of ensuring the implementation of the Equality Act, a broad anti-discrimination law passed by parliament in 1987, published a report that declared the quota practice in primary teacher admissions as discriminatory against female applicants (Tasa-Arvovaltuutetu, 1987). Since its lifting, politicians and the general public have repeatedly argued for the quota's reinstatement, motivated by the fact that boys are increasingly lagging behind academically and that a growing number of children are being raised by single mothers and lack a father figure (Etelä Suomen Sanomat, 1988; Liiten, 2012).

While we cannot directly observe how the quota and its lifting affected the gender composition of admits as data on university applicants at that time were not recorded, Figure 1 shows a proxy measure based on obtained degrees to illustrate how the share of men among cohorts of incoming primary school teachers changed over time. We plot the share of men among those who ever obtain a degree in primary teacher education against the year of their last attempt at the matriculation exam that qualifies students for university studies. Since students can repeat the exam multiple times if they want to increase their score, the year in which they last took the exam serves as the closest proxy for when they start their university studies: Anyone taking the exam in and after 1989 will with certainty have studied under the non-quota application regime. While the measure presented here only includes students who graduate with a primary teaching degree, we can still observe a sharp and sudden drop of about 15 percentage points in the share of male teachers among matriculation exam cohorts.[4]

Since the quota did not only change the gender composition of incoming teachers but also advantaged academically lower scoring men, Figure 2 plots future teachers' national percentile rank in the matriculation exam for the first attempt of the exam, against the last year in which they ever took this exam. We use the average national percentile score across all subjects for the *first* time that future teachers took the exam in order to get at a measure of inherent ability that is not influenced by repeated test taking, but we plot this percentile rank against the date of their *last* exam to most closely approximate the point of entry to university studies.[5] While the quota was in place, men on average scored about 10 percentile points lower. Once the quota was lifted, the average score among male teachers increased a bit, consistent with universities no longer admitting the worst scoring male applicants. We will return both to the changes in gender composition and academic scores more formally in section 6.1.

Teacher gender in our setting is correlated with a bundle of other characteristics that may matter for teaching. Table 1 presents summary statistics on male and female teachers who are active in the profession before the lifting of the quota (i.e. before 1994 as the year in which

[4] Notice that the gradual drop before 1989 is consistent with a setting in which students apply multiple years in a row. We can only assess whether the quota was binding in this indirect way. While the lifting of the quota was widely discussed in policy and media reports at the time, we have found no documentary evidence that either application numbers or the composition of applicants would have drastically changed with the lifting of the quota.

[5] Data on exam points for the year 1990 are missing, so that we cannot calculate the national distribution of scores according to percentiles for that year. Note that if the worst scorers are those that repeat the exam, this will bias the average scores in the years after 1990 upwards (the worst performers will not be counted in these averages as their scores are missing). We address this issue by also examining average grades, for which we have data in 1990. We can show that the pattern is very similar and upward bias in the percentile scores after 1990 should be small.

the first non-quota cohort graduates from teacher studies), and thereafter. We can observe that male teachers are slightly more likely to come from rural areas and to live in their region and municipality of birth when compared to female teachers. Regarding educational trajectories, there is no difference in having obtained a high school degree and being a certified teacher, but male teachers on average are a bit older both when finishing high school and when being awarded their teaching degree. Statistics on the matriculation exam show no difference in having passed the exam, but again illustrate that male teachers perform much lower, even when considering the best exam take for repeated attempts. High school students have some flexibility to choose either mathematics or a combination of other natural and social sciences ("Reaali") in the matriculation exam. Male teachers are about 9 percentage points more likely to have taken the mathematics exam compared to female teachers, and 10 percentage points more likely to have chosen advanced level mathematics rather than the basic level exam.[6]

2.3 Related literature

We contribute to two large branches of literature. First, we contribute to a large literature that has examined questions regarding the impacts of affirmative action policies on representation, quality and potential mismatch of marginal candidates. Studies of affirmative action policies have largely used different policies in the context of university admissions, most of them in the United States. Empirical evidence on quota policies, a subset of affirmative action policies which prescribe a fixed number or share of slots for a particular group, has to date almost exclusively examined two specific settings, both regarding female representation: quotas for board rooms in firms, and quotas in the context of political leadership.

A large body of work has documented that affirmative action policies succeed in raising representation for their intended beneficiaries from underrepresented minority (URM) groups (Bertrand et al., 2010; Yagan, 2016; Bleemer, 2021; Bleemer, 2020), but it has been a more contested issue whether lower levels of academic preparedness result in a mismatch of URM candidates to an environment that may be too demanding (Arcidiacono and Lovenheim, 2016). Recent quasi-experimental work on affirmative action policies at state flagship colleges in the United States suggests that these measures do benefit their target group as intended: higher likelihoods of attending and graduating from selective colleges result in earnings gains for at least some URM groups (Bleemer, 2020). At the same time, marginal students who are crowded out by these policies do not appear to suffer negative effects, rendering such policies both equity and efficiency enhancing on the margin (Bleemer, 2020; Black et al., 2020; Bleemer, 2021). In a similar vein, studies on quotas for female politicians have documented that 'mediocre men' are replaced with more highly qualified women in Swedish and Italian politics (Besley et al., 2017; Baltrunaite et al., 2014).[7]

[6] Each exam field and level of difficulty is graded on a curve within that group.

[7] In the context of the Norwegian board quotas, Bertrand et al. (2018) document that female board members hired due to the quota are more qualified relative to the average female board member before the quota took effect, but have lower levels of experience relative to the average male board room member.

Questions on whether affirmative action policies affect production or proxies for quality of output have only been examined in the context of quotas. Survey-based evidence from a quota rule for female village heads in India documents that female leaders invest more in public goods favored by female voters, but these findings do not necessarily imply that female quota politicians are more efficient in delivering public goods overall (Chattopadhyay and Duflo, 2004).[8] Evidence on how quotas for board rooms affect output related measures for firms has been inconclusive: The most frequently studied setting is Norway's introduction of a 40% quota for female board members on publicly listed companies in the mid-2000s. Studies examining the quota's effect on measures of firm performance arrive at contradictory results that depend on the chosen counterfactual. While some studies document a negative effect of the reform on Tobin's Q (Ahern and Dittmar, 2012) and profits (Matsa and Miller, 2013), others find neutral effects on stock market returns (Eckbo et al., 2021; see also Bertrand et al., 2018 for a more extensive review). Similarly, the direction of effects for analogous board quotas introduced in other European countries do not indicate a clear pattern.[9] A recent study by Cortés et al. (2021) documents negative effects on performance of a policy that required firms to hire native workers in Saudi Arabia.

In this paper, we can go one step further with respect to previous work on affirmative action and quota policies and cleanly document that preferential access to teacher studies does not just increase men's representation, but that their contributions to production in their industry lead to sizeable efficiency gains. We focus on a setting in which the entire composition of a broad and socially relevant occupation category is shifted by a quota rule and measures of output are readily observable in the long term. We further develop a conceptual framework to guide the analysis of how these efficiency gains may materialize despite male candidates' lower average quality in terms of academic achievement upon admission.

Second, we contribute to research that has been interested in examining the effects of teacher characteristics, and specifically teacher gender, on pupil outcomes. A large body of work has documented that measures of teacher value added, rather than certification or test scores (Kane et al., 2008; see Hanushek and Rivkin (2006) for a review), drive variation in pupil outcomes (Chetty et al., 2014; Rivkin et al., 2005). Several studies have documented that exposure to a same-gender or same-race teacher can improve academic performance (Dee, 2004; Lim and Meer, 2017, Carrell et al., 2010), and suggest that effects may be driven by role model effects as pupils become more academically motivated when paired with a teacher they identify with in

[8] Chattopadhyay and Duflo (2004) specifically highlight that it is through the characteristic of being female herself that the leader's preferences are more closely aligned with female constituents, thus moving the status quo to more closely reflect preferences of the median voter. Related studies are Beaman et al. (2012) and Beaman et al. (2009) who expand on the same setting and document that voters' gender role attitudes regarding female leaders change in villages in which the quota rule assigned female chiefs twice in a row. These changes in attitudes drive increased representation of female politicians in the longer term (Beaman et al., 2009) and a reduction in gender gaps for adolescents' aspirations and educational attainment, consistent with a role model channel (Beaman et al., 2012).

[9] Results from Italy show no effect of a female quota on performance, but a positive reaction of markets to the quota-induced turnover of boards (Ferrari et al., 2021), while Comi et al. (2020) document opposite effects of gender quotas in Belgium and France (negative) versus Spain and Italy (positive).

terms of gender or race (Dee, 2005; Sansone, 2017).[10] An active literature has studied such role model effects also with respect to short and medium-run course enrollment in mathematically oriented fields for women (Porter and Serra, 2019; Breda et al., 2018). These studies tend to document effects for specific underrepresented sub-groups, i.e. high achieving women attending a science track. Riise et al. (2020) show that exposure to a female GP in childhood raises women's attachment to natural science fields in the long run. Due to the level of detail and comprehensiveness of the Finnish register data, we can follow subjects throughout a substantial portion of their young adult life, and analyze their specialization choices in the long run. We add to prior studies by being able to observe a quota-induced shift in the gender composition of the entire occupation category of primary school teachers, thus being able to examine role-model effects in the light of quota policies.

3 Conceptual framework

This section develops a conceptual framework assessing the trade-offs between teacher quality and gender representation. In our setting, the admissions office is the decision maker who would like to maximize future teacher ability. However, true teacher ability is unobserved by the admissions office so that it selects candidates based on academic scores.[11] We outline how a quota's effect on total candidate quality depends on the interplay between scores as the selection criterion and teacher ability: We first examine a benchmark case in which scores fully reflect teaching ability, such that a quota is costly. We then proceed to examine two cases in which the mapping of scores into ability is imperfect, such that a quota has the potential to be efficiency enhancing: We discuss a case where the mapping of scores is group-specific, and a case in which there exist complementarities in production between teachers of different genders. We summarize these theoretical points by deriving testable predictions for the empirical part.

3.1 Set-up

Consider an admissions office that seeks to select a fixed mass of candidates C from a pool of applicants by considering their academic scores, s. Candidates belong to one of two groups g, with $g \in \{M, F\}$ for Male and Female, and are heterogeneous with respect to their score, s. Scores are defined on $[\underline{s}, \bar{s}]$, the density of scores $h_g(s)$ is twice continuously differentiable, and we assume full support, $h_g(s) > 0$ on the relevant interval. We assume that the distribution of scores of the M group is first order stochastically dominated by the distribution of F scores to

[10] Antecol et al. (2015) show that random exposure to a female primary school teacher with weak math background lowers girls' (but not boys') math performance, indicating that same-gender identity can also lead to undesired outcomes.

[11] Without loss of generality, we make two simplifications in the model relative to the actual selection process in our setting: We define academic scores as the combination of test scores in the matriculation exam plus any additional points received for extra curricular activities, as the matriculation exam receives a large weight in this calculation. Further, we only model the first step in the selection process where candidates are selected on scores alone. While universities would have some leeway of how many candidates to invite to the second step, we think that this is reasonable as the empirical evidence points to the quota in the first step driving final admissions.

reflect the empirical fact that male applicants tend to have lower scores in our setting:

$$H_F(s) \leq H_M(s)$$

for all s, with strict inequality whenever $H_M > 0$ and $H_F < 1$. The mass of candidates of each group that is admitted above a cutoff score s^* is given by $M = \int_{s^*}^{\bar{s}} h_M(s)ds = 1 - H_M(s^*)$ and $F = \int_{s^*}^{\bar{s}} h_F(s)ds = 1 - H_F(s^*)$.

We further introduce a measure of candidates' teaching ability, a, that we will return to in more detail in iterations of the model below. We define a welfare function W as the total teaching ability of admitted candidates for a given ability threshold, a^*:

$$W(a^*) = \int_{a^*}^{\bar{a}} a\, h_M(a)da + \int_{a^*}^{\bar{a}} a\, h_F(a)da \tag{1}$$

In our setting, the admissions office's aim is to pick the highest ability teachers, but since it can not observe a or is perhaps restricted to only consider s, it has to rely on candidates' scores as a proxy for ability.

Admissions without quota: The admissions office picks a threshold score s_g^* above which all candidates are admitted subject to a capacity constraint:[12]

$$\max_{s_M, s_F} \int_{s_M}^{\bar{s}} s\, h_M(s)ds + \int_{s_F}^{\bar{s}} s\, h_F(s)ds \tag{2}$$

$$s.t. \;\; M + F = C$$

When maximizing total scores of admits, the optimal cutoff s_g^* is the same for both groups: $s_M^* = s_F^* = s^*$, resulting in the mass of admitted women outnumbering admitted men due to first order stochastic dominance.

Admissions with quota: Suppose we now introduce a quota rule that adds an additional constraint to the admissions office's optimization by requiring that at least mass q of students from group M be admitted. The admissions office then solves the problem in equation 2 with the additional constraint:

$$s.t. \;\; M \geq q \tag{3}$$

[12]The admission office's problem here is similar to Chan and Eyster (2003) who make the theoretical point that forbidding universities whose utility function contains an exogenous preference parameter for diversity from using race as a selection criteria will result in lower quality of admits. This results from the optimal admissions rule under race-blind admissions consisting in randomly selecting applicants above a lower cut-off score, whereas color-sighted admissions select the best scoring candidates within each group. Ellison and Pathak (2021) bring the reasoning of Chan and Eyster (2003) to the data to evaluate the efficiency of a place-based affirmative action rule in two Chicago Public Schools. Instead of assuming a preference parameter for diversity, their model defines students' outcomes as a trade off between an optimal level of school diversity and academic match. Our set-up does not assume a taste parameter for diversity.

When the quota constraint is binding, the admission cut-offs between groups diverge and the admissions office chooses optimal cutoffs that are group-specific:

$$s_M^*(q) = s_F^*(q) - \delta \tag{4}$$

where δ is the Lagrange multiplier of the quota constraint and indicates the shadow price of admitting a male candidate at the margin. The following iterations of the model briefly discuss how total teaching ability evaluated by W changes for different mappings of scores s into teaching ability a when the admissions office picks optimal cut-off scores, relative to admissions under full information.

3.2 Candidate choice when scores fully reflect ability

We start by evaluating a benchmark case: Scores fully reflect teaching ability, so that

$$a = s$$

In this case, requiring the admissions office via a quota to admit more men relative to the unconstrained optimization in equation 2 will result in total teaching ability of admits being lower under a quota: $W(s^*(q)) < W(s^*)$. This illustrates one of the main concerns frequently brought forward against quota rules: Requiring the admissions office to forgo its preferred allocation of slots will lower total candidate quality, as the ability of additional men admitted due to the quota is less than the ability of women who must be rejected to satisfy the quota:

$$\int_{s_M^*(q)}^{s^*} s\, h_M(s) ds < \int_{s^*}^{s_F^*(q)} s\, h_F(s) ds \tag{5}$$

3.3 Group-specific mapping of scores into ability

Consider now a case in which the association between scores and teaching ability is group-specific:

$$a = s + \alpha \mathbb{1}_M$$

with $\alpha > 0$, $\mathbb{1}_M$ an indicator function that denotes membership in group M, and the density ρ such that $\rho(s + \alpha \mathbb{1}_M) = h(s)$. For any given teaching ability level a, the score of members of the M group is reduced by a penalty parameter α.[13] In such a world, switching out a marginal female with a marginal male teacher who has the same score increases total teacher ability.

Full information: When evaluating total teaching ability of admits under full information, W takes into account men's true ability. Substituting for a with the above relation, the optimization problem when fully observing teacher ability becomes:

[13] Note that richer models deliver a penalty parameter of this form. For example, when scores are a differentially noisy signal of ability for different groups, the conditional expectation of ability given score may take this form (Phelps, 1972; Aigner and Cain, 1977).

$$\max_{s_M, s_F} \int_{s_M}^{\bar{s}} (s+\alpha)\rho_M(s+\alpha)ds + \int_{s_F}^{\bar{s}} s\rho_F(s)ds \tag{6}$$

$$s.t.\ \ M + F = C$$

The first order conditions under full information (FI) imply that the optimal cutoff scores in this case are such that:

$$s_M^{FI} = s_F^{FI} - \alpha$$

The marginal male admit here has a strictly lower score than the marginal female admit and the optimal allocation of slots fully takes into account that M group members' scores suffer a test score penalty α for any given ability level a.

Admissions Office: In contrast, the admissions office continues to select candidates based on scores alone (equation 2): $s_M^* = s_F^* = s^*$. In this case, it is the lack of differential score thresholds that is costly as the admissions office does not take into account men's score penalty, and $W(s^*) < W(s^{FI})$.

With a quota rule, the admissions office chooses $s_M^*(q) = s_F^*(q) - \delta$. The wedge between marginal male and female admission scores grows with the mass of male candidates required by the quota, and achieves the optimum under full information when the quota is set such that $\delta = \alpha$. This illustrates that the admissions office deviates further from the optimum under full information both when a quota under-, as well as when it over-corrects for differences in scores: Total teaching ability will steadily decline with the mass of men required by the quota rule once $\delta > \alpha$, and will eventually yield an outcome worse than $W(s^*)$.

3.4 Candidate choice with complementarities

Even when scores fully reflect individual teacher ability, the existence of complementarities between M and F teachers can motivate differential treatment by group status. If the gains in total teacher ability from complementarities outweigh the costs in terms of individual scores, switching out marginal female with marginal male teachers will increase overall teacher quality, but with decreasing marginal returns. This could be the case if, for example, pupils are more motivated when exposed to same-gender teachers, or when diverse teams of male and female teachers perform better when designing school curricula.

We return to the assumption in the benchmark case that scores perfectly map into ability, $s = a$. We introduce complementarity in teaching ability with a production function V that exhibits positive cross-derivatives, and treats M and F candidates symmetrically:[14]

$$V = \sqrt{L_M L_F}$$

[14] We focus here on homogeneity of degree 1, which implies that double the amount of teachers can teach double the amount of pupils, but any degree of homogeneity larger than zero yields the same conclusions

with $L_M = \int_{s_M}^{\bar{s}} s\, h_M(s) ds$ and $L_F = \int_{s_F}^{\bar{s}} s\, h_F(s) ds$ as total teacher ability for group M and F respectively. Under full information of V, the optimum equates the relative test score thresholds to the relative total ability of admitted candidates:

$$\frac{s_M^{FI}}{s_F^{FI}} = \frac{L_M}{L_F}$$

For a common threshold score $s_M^{FI} = s_F^{FI}$, it has to be that $L_M = L_F$, which is inconsistent with the assumption that $H_F(s)$ first order stochastically dominates $H_M(s)$ with strict inequality. The threshold for M must be lower, so that in the optimal allocation, $s_M^{FI} < s_F^{FI}$.

We therefore have three potential descriptions of the relationship between test scores, output, and the effect of quotas. If test scores perfectly reflect future teacher ability, quotas will reduce output in the absence of complementarities. Our empirical set-up directly serves as a test of whether the quota was efficieny-enhancing by examining output in terms of pupils' outcomes. If test scores imperfectly reflect teacher ability, so that men of a given test score tend to have higher ability than women, a quota may improve output so long as it is not too stringent. Complementarities may cause greater output from male teachers with lower test scores even if all components of ability are reflected in scores, but will do so with decreasing marginal returns.

4 Data and sample

Our main data source is register data maintained by Statistics Finland which span the years 1988 - 2018, and contain detailed yearly information on all residents of Finland. We compile two main data sets that correspond to the respective parts of the analysis. For primary school teachers, we construct a panel of active primary school teachers from 1990 - 2000 for all individuals whose occupation at any point in time between 1990 - 2005 is classified as a primary school teacher by Statistics Finland's occupation classification system in the employment register. Since occupation categories are first available in 1990 and are not reported in every year, we use a combination of workplace, industry, salary, degree and career information to infer active teacher status in any given year (FOLK employment, basic, and degree). We can match teachers' matriculation exams scores and dates for all cohorts born after 1952 (YTL moduuli), but we unfortunately do not observe university enrollment or study progress for teachers as these registers were not maintained at the time.

For the pupils, we observe the universe of children living in Finland who turn seven years old (and therefore start school in that calendar year) between 1988-2000. We assign them to a municipality (and teacher gender composition during grades 1-6) based on their place of residence in the year in which they start school. We further match pupils to their parents which allows us to observe a rich host of variables related to families' socio-demographic characteristics at age seven (FOLK family). We merge the pupil cohorts to registers on upper secondary education applications (EDU-THYR), registers that document enrollment in post-compulsory education

(EDU-OPISK, only for school starting cohorts from 1990 onwards), as well as the population, degree and employment registers that are available on a yearly basis after age 16.

We measure all of the treatment variables at the municipal level as data to link pupils and teachers to classrooms or schools does not exist. As our main goal is to estimate the impact of a quota per se, and not the impact of having a teacher of a particular gender, aggregating the data to a level higher than the classroom is consistent with both the research question and a setting in which collegial collaboration is widely practiced. The median population size among the 461 municipalities in our sample is 5000 inhabitants in 1990, with a median of 73 children by cohort in primary school. Appendix Figure A13 shows the CDF of the number of primary and middle schools combined across municipalities in the year 2005, as earlier or more detailed data was unobtainable. About 20% of municipalities have fewer than 5 primary or middle schools.

To comply with data disclosure regulations by Statistics Finland, we exclude municipalities that contain fewer than three teacher observations in a given year from our analysis. Once we move to a pupil level panel, we restrict the sample to municipality*year cells for which we are able to observe at least six teacher observations (i.e. the teaching staff for grades 1-6 of one school) to deal with noise arising from overlapping cohorts.

5 Empirical strategy

We want to study whether and how output is affected when the gender composition of teachers changes via a quota. Lifting the quota at the point of university admissions will impact the gender composition among active primary school teachers only gradually over time, but the changes in the flow of incoming teachers are sharp and immediate. In the estimation strategy, we therefore use shocks to the demand for new teachers in municipalities that arise from idiosynchratic local teacher retirement. Since teacher retirement could respond endogenously to the policy reform itself, we only use variation from teachers reaching the union mandated retirement age of 60. We use the term "retirement" exclusively to refer to teachers turning 60 throughout the paper.

Ideally, we would like to observe yearly pupil level outcomes and classroom specific links of pupils and teachers in order to differentiate between direct and indirect exposure effects. Unfortunately, such data does not exist. An ideal experiment, that takes the aggregation level of our data as given, would consist in randomly removing some teachers from municipalities, and deciding with a coin flip whether replacement teachers will be drawn either from a pool of only male quota teachers, or from a pool with only marginal female teachers.

Our DiD-IV estimation strategy closely approximates this experiment, taking into account that changes in quota teachers materialize via the inflow of rookie teachers and that we cannot observe quota male and marginal female teachers in the data. Municipalities in our setting are randomly assigned quota men – and thus more male teachers in general – via the *timing* of their open positions arising from teacher retirement. Our IV strategy estimates a local average treatment effect for complier municipalities: Variation stems from those municipalities that via the timing of retirements are induced to hire more vs less quota men among their teachers. While this notion matters for assessing the external validity of results, we think that our estimates get us close to the policy relevant parameter of interest: What happens when we change the composition of an occupation via a quota that operates through the inflow of incoming candidates?

Figure 3 outlines the timeline of our reform: The primary school teacher students who enter university before 1989 are selected via the quota rule. As the official time to complete the degree is five years, the quota and non-quota cohorts of new teachers will leave university around the year 1994 and will be hired by municipalities for their local schools. If municipalities have open positions during the time when quota cohorts enter the teacher market, they will be more likely to hire candidates from a pool with relatively more male rookie teachers compared to municipalities that have to fill open positions once new teacher cohorts selected without the quota are entering the teacher market.

5.1 Municipal level: Changes in teacher composition

We first document that local retirement interacted with the timing of abolishing the quota changes the local gender composition of teachers. Consider the following fixed effects specification:

$$\textit{share male}_{mt} = \pi_0 + \pi_1 \,\textit{total share 60}_{mt} + \pi_2 \,\mathbb{1}_{t=post}\, \textit{total share 60}_{mt} + X_{mt}\delta + \eta_{rt} + \gamma_{mp} + \zeta_{mt} \quad (7)$$

with $\textit{share male}_{mt}$ the share of male teachers in municipality m in a given year t, and $\textit{total share 60}_{mt}$ the total teacher retirements in a municipality up to that point in time.[15] The indicator function $\mathbb{1}_{t=post}$ switches on once non-quota teacher cohorts graduate and start entering the teacher market in the year 1994. The coefficient of interest, π_2, measures how additional retirements in the post-quota period affect the share of local male teachers *relative* to when the quota was still in place. We add region-by-year fixed effects η_{rt} to control for time-varying shocks whose impacts may vary regionally, with a total of 19 regions comprising on average 24 municipalities. In addition, we can also add controls for time-varying municipal characteristics X_{mt}. The municipality-by-period fixed effects γ_{mp} 'reset' the measure of total retirements to zero once the post period starts to separately estimate how retirements affect the local share of male teachers in the post period.[16]

[15] The fixed effects specification of equation 7 uses the stock of the dependent variable (the share of male teachers) and the independent variable (the cumulative share retiring teachers over time). The corresponding first difference equation uses flows on both sides of the equation by regressing the year-on-year changes in the share of male teachers within municipality on the share retiring teachers in each year, dropping the municipal fixed effects:

$$\Delta \textit{share male}_{mt} = \pi_0 + \pi_1 \,\textit{share 60}_{mt} + \pi_2 \,\mathbb{1}_{t=post}\, \textit{share 60}_{mt} + X_{mt}\delta + \eta_{rt} + \zeta_{mt} \quad (8)$$

We report first stage results for both equations, and use equation 8 when thinking in flows is more intuitive for some robustness checks on hiring patterns.

[16] The reset is necessary so as to properly net out any effect of the quota-period retirements from the post-quota estimate. The effect of retirements on the gender composition in the post should be independent of how much retirement the municipality faced in the quota period.

5.2 Pupil level: Does the quota shift in teacher gender affect outcomes?

Structural equation: Our main equation of interest is the following structural equation:

$$y_{im,t+x} = \beta_0 + \beta_1 \overline{\textit{share male}}_{mt} + \beta_2 \overline{\textit{total share 60}}_{mt} + X_i\delta + \gamma_m + \eta_{rt} + u_{imt} \quad (9)$$

with $y_{im,t+x}$ the outcome of interest at time $t + x$ for pupil i who at age seven lived in municipality m, and X_i individual level controls also measured at age seven, such as parents' socio-economic status. We add municipal fixed effects γ_m, as well as region-by-cohort fixed effects, η_{rt}. We are interested in how increasing the average share of male teachers via the quota affects pupils' outcomes, with $\overline{\textit{share male}}_{mt}$ the average of the share of male teachers across the years we observe pupils in primary school.[17]

Our empirical strategy isolates variation in the share of male quota teachers from gender changes in the inflow of recently graduated teachers that are caused by retirements. Rookie teachers may differ from older teachers along various dimensions: they have less experience, but they might also be differentially motivated to teach. We control for the overall potential share of rookie teachers a pupil is exposed to via retirements by controlling for the average aggregate share of teacher retirements during a pupils' time in primary school, $\overline{\textit{total share 60}}_{mt}$. This measure is the average across all cumulative retirements a pupil experiences from grade 1-6, and we discuss it in more detail below. Note that it is possible that schools' hiring decisions, and thus the impact of being exposed to retirements, may change due to the quota.[18] As such, our estimates measure the total effect of the policy that could include differential responses to retirement shocks. We elaborate on this in more detail in the robustness section 8.1, and do not find evidence that schools would change their hiring practices over time.

First stage: We instrument for $\overline{\textit{share male}}_{mt}$ with the following first stage equation on the pupil level that closely mimics the municipal level first stage in equation 7. Since every time period t corresponds to a particular cohort starting school at time t, we will refer to t as a cohort identifier in the following:

$$\overline{\textit{share male}}_{mt} = \pi_0 + \pi_1 \overline{\textit{total share 60}}_{mt} + \pi_2 \, \mathbb{1}_{t=post} \overline{\textit{total share 60}}_{mt} + X_i\delta + \gamma_m + \eta_{rt} + \epsilon_{imt} \quad (10)$$

Variation in treatment intensity arises from how much teacher retirement different cohorts of pupils across different municipalities experience in the post-quota relative to the quota period. The coefficient of interest, π_2, measures how the share of male teachers a pupil experiences is affected by retirements in the post-quota relative to the quota period. By measuring the differential impact of retirements, we compare the causal effect of being exposed to new teachers against the causal effect of being exposed to new teachers with a changed gender composition due

[17] Our pupil panels spans 13 cohorts that are starting school in the years 1988-2000, and thus experience teachers who we can observe from 1990 - 2000. For some cohorts of pupils, we observe the teacher composition for each year that pupils are in school, while for others, we only know it for their starting or ending years. Appendix Figure A12 depicts the cohorts over time observed in our data.

[18] For example, if schools had hired fewer rookie teachers for each retiring colleague in the quota period because they did not want to select quota men.

to the lifting of the quota. In the structural equation, β_1 then measures the causal effect of being exposed to relatively more male teachers via incoming quota men. This relative comparison addresses exclusion restriction concerns that retirement triggered increases in rookie teachers matter for pupil outcomes.

We measure pupils exposure to retirements during their time in primary school, $\overline{\textit{total share 60}}_{mt}$, by taking the average over the cumulative retirements a pupil is exposed to during their six grade levels g in primary school:

$$\overline{\textit{total share 60}}_{mt} = \frac{1}{6}\sum_{g=1}^{6} R_{mtg} \tag{11}$$

with $R_g = \textit{share 60}_g + R_{g-1}$ and $R_1 = \textit{share 60}_{-2} + \textit{share 60}_{-1} + \textit{share 60}_1$

For example, the retirement measure for grade 6 (R_6) adds retirements that occur just before a pupil enters grade 6 ($\textit{share 60}_6$) to all retirements the pupil has experienced up to this point: $R_6 = \textit{share 60}_6 + R_5$. R_1 considers all retirements up to two years before a pupil starts school. We construct $\overline{\textit{total share 60}}_{mt}$ in this way to reflect the fact that retirements that happen later in the pupils' school career will have an impact on the teacher composition for relatively fewer years compared to retirements when pupils start school: Retirements that occur before a pupil enters grade 1 have the potential to change the gender composition, and thus the average share of male teachers, for all grades a pupil spends in primary school. In contrast, any retirements occurring just before a pupil enters grade 6 will affect the teacher composition only in their last year in school. In the empirical analysis, we report grade level results for the first stage that directly motivate the construction of this measure.

5.3 Discussion of identifying assumptions

We revisit explicit and implicit identifying assumptions of our setting in more depth. To start with, our identification strategy needs to satisfy the two main IV assumptions. Relevance requires that teacher retirements in the post-quota period decisively impact the local share of male teachers, which we can assess directly in the first stage regressions.

The exclusion restriction, briefly touched on in the above, warrants more discussion: We require that teacher retirements affect pupils' outcomes only via changes in the share of male teachers, and thus changes in male quota teachers. However, retirements themselves, by triggering teacher turnover, may have a direct effect on pupils. We tackle this by measuring relative changes in outcomes between cohorts that experience similar exposure to retirements, but with different timing. The underlying assumptions here are twofold: First, we need to assume that there are no other policy changes that happen simultaneously with the quota that have effects on students *via the channel of retirements.* To the best of our knowledge, there are no such policies.[19] Secondly, we assume that exiting patterns and hiring practices to replace retiring

[19] As an example, consider a policy where retired teachers would regularly come in to schools to mentor new

teachers do not differentially change as a response to the quota.[20] We can test for such patterns in Section 8.1 and do not find evidence for differential changes in the post period.

Implicit in our empirical design is the further assumption that the local timing of retirements is idiosyncratic, and therefore uncorrelated with any other shocks that could affect pupil outcomes.[21] We address such concerns by only using variation arising from teachers turning 60 (instead of actual exits), by controlling for a rich host of pupils' socio-economic characteristics at age 7, and by including region-by-cohort fixed effects. As such, we are only comparing cohorts in municipalities within the same region and year, with the notion that relevant economic shocks (in the past and currently) will similarly affect neighboring places.

Finally, while our regressions are measuring the effect of having more male quota teachers, we see teacher gender not just as a biological distinction, but as something that proxies for a bundle of characteristics that may differentiate quota male and marginal female teachers.

6 Main results

6.1 Municipal level: Effects on teacher gender composition

We start by documenting the effects on local teacher gender composition at the municipal level after the quota was lifted. We first examine teacher exit patterns: Figure 4 plots the exit probability by age for all primary school teachers in our sample by teacher age, reporting the probability of a teacher not teaching at a given age, conditional on having been a teacher in the previous year. There are some exits around the time when teachers become mothers, but quite notably a large spike exactly at the union mandated retirement age of 60. In our estimation, we are only using variation from teacher exits that is due to teachers complying with the union mandated retirement age.[22]

To illustrate the intuition of the first stage using the raw data, Figure 5 graphically illustrates the relationship between teacher retirement in a municipality (on the horizontal axis) and changes in the share of male teachers by separately plotting the period in which quota cohorts enter the teacher market (years 1991-93) and a period of similar length in the post-quota period (1994-96). We can see that teacher retirement has a small positive effect on the local share of male teachers in the quota years. In the post-quota period, higher shares of teachers retiring are associated

colleagues in the quota period, thus making rookie teachers 'better teachers', but that this practice is not continued in the post period.

[20]This includes the monotonicity assumption that rules out defiers in a LATE framework. In our case, these would be municipalities that would not want to hire male teachers while the quota is in place when facing retirements, but start hiring differentially more male teachers for retiring teachers in the post period.

[21]Consider, for example, an economic boom in the past that led to a big hiring spree in teachers, who are turning 60 within our time window. If prior economic shocks are correlated with current economic characteristics that affect pupil outcomes, such a pattern would be a concern.

[22]We also examine teachers' likelihood of changing jobs across municipalities in Appendix Figure A1. Less than 1% of teachers in the age bracket above 55 are changing the location of where they teach across all years of our panel.

with substantial local drops in the share of male teachers.[23]

Figure 6 formalizes this intuition by running the first stage equation 7 as an event study, estimating separate coefficients year-by-year, relative to 1993 as the baseline year. We can see that – similar to Figure 5 – teacher retirements in the years in which the quota was still in place do not differentially have an effect on the local share of male teachers relative to the year 1993, while retirements in the post period lead to a sizeable drop of about 20 percentage points. Table 2 summarizes this result for both the first difference and fixed effects specifications, by bundling years into a quota and post-quota period. Results are quantitatively similar across specifications. While retirements in the pre-period have a small positive, but insignificant effect on the local share of male teachers, the coefficient of interest on retirements in the post-quota period consistently reports large drops. We assess potential negative weights arising from two-way fixed effects estimation in our setting in section 8.2 following De Chaisemartin and d'Haultfoeuille (2020).

The reported coefficients correspond to measuring what would happen if all teachers in a municipality were to retire in the post-quota period: In this scenario, the local share of male teachers would drop by about 17-22 percentage points. These magnitudes match the drop in incoming male teachers reported by the literature and observed in teacher graduates in Figure 1. We can re-scale this coefficient to reflect a more realistic retirement pattern: If 10% of local teachers reach age 60 in the post-quota period, this translates into a 2 percentage point drop in the share of male teachers, which corresponds to a 5.5% decrease over the mean in the baseline period.

6.2 Municipal level: Effects on teachers' academic ability

While the quota targeted the gender composition of incoming primary school teachers, it simultaneously affected overall academic ability among teachers by giving preferential access to men with lower academic scores on average. While we only observe matriculation exam scores for teachers born after 1952, we can report a first stage estimating the effect of teacher retirement on average academic scores of teachers at the municipal level using these cohorts.[24] We restrict the sample to those municipalities for which we observe at least one teacher with a score in the baseline period. In Table 3, we report the results across different fields of the matriculation exam, using teachers' score in their first attempt at the exam. Albeit quite noisily estimated, the coefficient for average scores in column 1 shows that retirements in the post period lead to an increase of about 1.25 percentile scores in the local teacher body, relative to the quota period. This magnitude is consistent with replacing approximately 20% of teachers with an on average 7-8 percentile point higher test score in the post period (see Figure 2).[25] Coefficients are of

[23] Note that since our teacher panel spans 1990-2000, the first time we can calculate the share of teachers turning 60 that is relevant for re-hiring for the upcoming academic year is for 1991 (i.e. the 1991/92 academic school year)

[24] We observe scores for roughly 59% of the total teacher sample. For all specifications in which we need to restrict the sample due to data availability, we also run the first stage regression to ensure comparability. Those coefficients are reported in Appendix Table A1.

[25] The average decrease in test scores estimated for the quota period relative to the stock of teachers are also present in the raw data and represent a general downward trend in teacher test scores (Izadi, 2021).

similar magnitudes and precision when looking separately at different fields of the matriculation exam in columns 2 - 4.

Insofar as one might think that academic test scores translate into actual teacher quality or ability, we observe small increases in average local test scores that are noisily estimated as we do not observe score data for the entire sample of teachers. These small increases in scores accompany quite sizeable impacts on local teacher gender composition in the first stage. We next turn to examine how these changes affect pupils.

6.3 Pupils: Labor force and educational attachment at age 25

We next turn to our panel of pupils, spanning the cohorts that enter primary school between the years 1988 to 2000. We start by documenting the first stage relationship: are children that experience more teacher retirement post-quota exposed to fewer male quota teachers? Our outcome here is the average share of male teachers across all of the six years a pupil spends in primary school.

Figure 7a shows the first stage results if we regress the average share of male teachers on the share of retirements pupils experience just before they start each grade level, starting up to two years before they enter school and until grade six.[26] Figure 7a depicts coefficients separately for the quota period (grey) and the post-quota period (green), while Figure 7b shows the effect of retirements in the post-quota period *relative* to the quota period. Teacher retirements in the early years of students' primary school time have a large and significant impact on the average share of male teachers pupils experience during their time in primary school, with similar magnitudes as the municipal level first stage results. At higher grade levels, this effect gradually peters out. This pattern is as expected as it shows that retirements in early grades, which affect the teacher composition during the entire six years a pupil spends in primary school, matter more for explaining the exposure measure which averages the share of male teachers a pupil faces across their entire time in primary school. Similarly, retirements that happen just before a pupil enters grade six will only impact the share of male teachers for one year, and therefore contribute less to moving the average share of male teachers over all six years. This pattern, as explained in section 5.2, informs our construction of the instrument when bundling a pupil's exposure measure across all grades: we define a pupils' average total exposure to retirements as the cumulative share of teachers retiring in each grade level, which gives more weight to retirements that have the potential to change the teacher composition for a longer time.

In Table 4, columns 1-3 show results for the pupil level first stage. Due to the cumulative nature of the explanatory variable, we can interpret this coefficient as 'how much does the average share of male teachers change when there is full retirement before a pupil starts school'. Results are similar to the municipal level regressions: Pupils facing 10% of teachers retiring just before they start school are exposed to about 1.8 percentage points fewer male teachers.

[26] Appendix Figure A12 gives a detailed overview over which cohorts are exposed to quota years in which grade levels.

Turning to outcomes, we examine pupils' labor market attachment at age 25 as reported to Statistics Finland in the last week of the calendar year. As many youths are still studying at this age, but are classified as employed due to part time work, we combine the categories of being a student and being employed into one measure that reflects not sitting idle. For this age group, this metric is considered relevant to measure the propensity to successfully integrate into the labor market (Eurostat, 2021; OECD, 2021). Columns 7-9 in Table 4 report results: Being exposed to more quota teachers during primary school results in higher likelihood of being either employed or a student at age 25.

The coefficients report the effect size associated with an increase of male (quota) teachers from zero to all of the teaching staff being male (quota) teachers. To scale effect sizes to match a more realistic pattern, we can consider the impact of a 1 SD increase in the share of male (quota) teachers, which is around 0.065. For a 1 SD increase in the share of male (quota) teachers, pupils have a 0.03 increased likelihood of working or studying, which corresponds to a 4% increase over the mean. Translated into standardized effect sizes (see the bottom row of Table 5), exposure to 1 SD higher share of male teachers during primary school leads to a .09 SD increased attachment to the labor/education market.[27]

Figure 8a and the corresponding Table 5 report effects for our preferred specification (column 6 of Table 4) for the remaining, mutually exclusive labor market status categories. While we observe no effect on unemployment, pupils are somewhat less likely to be on a disability pension, and significantly less likely to be out of the labor force for other reasons. Figure 8b and Appendix Table A5 show results by gender. Instead of splitting the sample, we run our main specification (Equation 9) and include separate treatment effects for boys and girls while estimating controls and fixed effects jointly. Results are qualitatively similar to splitting the sample, but less noisily estimated.[28] Regarding labor market status, the coefficients by gender seem somewhat higher for boys, but are overall not statistically different from each other.

Next, we turn to examine educational attainment at age 25. After compulsory education, the Finnish education system has two tiers: vocational and academic. Standard three year vocational degrees offer training in occupation-specific skills. In addition to working towards the completion of a basic vocational degree, pupils may take academic high school coursework that qualifies them to study a broader range of subjects at higher education institutions and adds an extra year to their study time. Alternatively, or in addition, pupils can also take further specialization training, that expands and deepens occupation-specific skills.[29] While students

[27] We can also ask how many pupils in a school this corresponds to. An increase in the share of male teachers of 0.065 corresponds roughly to switching out 1 in 15 teachers from marginal female to quota male at a local school. As the average class size is 20 pupils, this place would have a total of 300 pupils, and therefore about 9 pupils switch their labor market status.

[28] Estimating heterogeneity by pupil gender requires taking a stance on how to think about common shocks that are absorbed by the municipality and cohort*region fixed effects. When estimating these fixed effects jointly, the underlying assumption is that we expect e.g. time-varying economic shocks to affect the choices of boys and girls to a similar extent. We also report split sample results for the main outcomes in Appendix X, and these do not differ from the joint estimation (*coming soon*).

[29] An example for a basic vocational degree is training to work in the vehicles sector which covers subjects from car sales to vehicle mechanics, while additional qualifications allow pupils to specialize e.g. in specific areas of

from the vocational track may qualify for specific fields of tertiary education at university, the typical study path for the tertiary level is at polytechnics. The academic path leads from a three year high school degree to a Bachelor's degree (3 years) and a Master's degree (2 years) at university. Appendix Figure A10 shows the organization of the Finnish education system in detail.

Figure 9a and Table 6 present results for educational attainment by examining the highest degree achieved at age 25 using mutually exclusive education categories. We can see a shift towards higher attainment both in vocationally oriented as well as in academic education paths. As such, we observe a shift away from remaining with compulsory education or a standard three year vocational degree only, towards a "vocational plus degree", defined as vocational degree holders with additional specialist or high school qualifications. A 1 SD increase in the share of male (quota) teachers makes pupils .09 SD more likely to shift towards such a higher skilled degree. Turning to academically oriented degrees, we similarly observe a shift away from high school degrees towards having completed a university bachelor level degree. When examining results by gender in Figure 9b and Appendix Table A6, the coefficients again do not differ systematically (with the exception of compulsory education). Overall, results with respect to labor market outcomes seem to be driven slightly more by boys, while some educational attainment categories are somewhat more pronounced for girls.

For the female pupils, we can also look at realized fertility up to age 26. In Finland, the average age at first birth increased from 27.2 in 1995 to 29 in 2016 and is close to the OECD average (OECD, 2019). Consistent with our finding that both men and women invest more in education and have a higher attachment to the labor force when exposed to more male quota teachers, we can document in Appendix Figure A2 and Appendix Table A11 that female pupils are less likely to have given birth by age 26, which is indicative of delaying fertility.[30]

7 Mechanisms

Where do these effects at age 25 stem from, and how would male quota teachers make a difference? This part explores in more detail the step-by-step educational decisions that lead to the observed differences when pupils are young adults. We then turn to examine if effects derive from male teachers inspiring pupils to pursue different educational fields. Lastly, we return to the predictions of our conceptual framework in an attempt to differentiate between a mechanism where either male quota teachers make teacher teams more productive via complementarities, or where male quota teachers are of overall higher teaching ability than the marginal female teachers they replace.

vehicle repair.

[30] Appendix Figure A3 and Table A12 show results for male pupils: There is a small increase in the likelihood of having a first child by age 26, but it is statistically not significant and economically small.

7.1 Intermediate outcomes: Applications to post-compulsory education

In order to better understand where the long run effects measured at age 25 originate, we next examine application choices to higher education options that take place after compulsory schooling at age 16.[31] After primary school (grades 1-6) and middle school (grades 7-9), pupils in Finland have the option to apply to upper secondary education, which typically takes three years to complete, is provided free of charge, and is divided into two tracks. In the vocational track, students acquire occupation-specific skills in a chosen field. In the academic track, pupils attend so called 'high schools' that prepare them to enter university, with pupils taking the national matriculation exam in their final year.[32] In grade 9, the final year of middle school, pupils apply for their desired institution, and in the case of the vocational track also their desired field. While further education is not mandatory after age 16, application rates for upper secondary degrees are very high with around 90% of a cohort applying in early spring of their final year in middle school. Raising completion rates of upper secondary education is a policy priority as a post-compulsory degree is deemed crucial for labor force attachment: Finns with only basic education are four times more likely to be out of the labor force, and have significantly lower employment rates in adulthood (Virtanen, 2016; Niemi et al., 2016). Prior research with Finnish data has shown that the choices and slot allocations in upper secondary education can be pivotal: With an RDD design, Virtanen (2016) shows that failing to obtain a preferred choice or a study slot at all results in a higher probability of drop out.

In the centrally organized application process, each pupil can submit up to five choices for institution (and field), and a student proposed ranked choice algorithm allocates available study slots.[33] Institutions rank applicants based on grades and other qualifications such as extracurriculars. As applications take place before pupils obtain their final grades that are used to allocate slots, and with the popularity of institutions and fields varying over the years, students face uncertainty over whether they are able to obtain a study slot. The number of available slots per degree is centrally regulated and about 4% of a cohort end up without a study slot in the fall after finishing middle school.[34]

Figure 10a and Table 8 report results on pupils' application timing and choices (Table 7 reports the first stage and main outcome for this section separately). Having more male quota teachers makes pupils more likely to apply directly in their final year of middle school at age 16, and less likely to postpone applying to up to five years later. We see no significant effect on never applying to upper secondary education. When considering the allocation of slots, we further observe that pupils are more likely to get one of their top two choices, while the effect

[31]Virtually everyone (99.7% of a cohort) successfully graduate from compulsory education (Virtanen (2016)).

[32]In principle, taking the matriculation exam and switching from a vocational track to the academic path is possible at any point, but rare in practice. See Appendix Figure A10 for more details on the Finnish education system.

[33]For an infinite number of choices, the algorithm would be strategy proof. Since students can only submit five choices, some may choose to enter a 'safe' option to make sure they get a study spot.

[34]See also Virtanen (2016) for an in-depth description of the allocation process of slots for upper secondary schooling.

of not getting any slot at all is more muted. Consistent with Virtanen (2016), we find that these patterns translate into higher enrollment rates in upper secondary education in general, and significantly so in the year in which students turn 16. Figure 10b and Appendix Table A7 report heterogeneity by pupil gender, with coefficients not statistically different for any outcome.

Why are pupils that are exposed to more male quota teachers more successful in obtaining their preferred choice? We can check whether pupils apply more cautiously, with their main choice between aiming for an academic high school degree or a vocational training. We report effects in this part directly by pupil gender, as for these measures results differ significantly and overall effects mask more intricate patterns. Appendix Table A8 reports results. Male pupils are more likely to include any vocational training option among their choices (column 2), while refraining from applying exclusively to high schools (column 3). For the girls, while not statistically significant, the effect goes in the opposite direction. Overall, boys seem to become somewhat more cautious in their applications, and girls more confident. When examining for which track options pupils obtain a slot in columns 4-6, we see that the margin for boys shifts from not obtaining a slot at all (column 4) towards getting a vocational spot (column 5), while the girls become more likely to obtain an academic spot (column 6) rather than a vocational one. Taken together, these results imply that having more male quota teachers makes pupils apply more in line with attainable options: Boys adjust their aspirations downwards, which prevents them from ending up without a slot at all, and girls correctly have high aspirations as they are more likely to get into academic high schools.

7.2 Education fields

We observe overall effects that – except for the adjustment of aspirations in applications to upper secondary education – do not differ by gender. These patterns are not consistent with a classical same-gender role model story of boys being academically more motivated when being exposed to a male teacher. However, male teachers may still set an important example of men working in an occupation that is female dominated. As such, they may inspire primarily male pupils to pursue a teaching-related field. On the other hand – and separate from a classical same-gender role model effect – male teachers could via their skills and motivation for particular fields make pupils more motivated to pursue different education fields. As we have seen in section 2.2, male teachers are on average more likely to have chosen advanced math as one of their matriculation exam fields, and may thus be more skilled or motivated to teach mathematics.

In order to investigate these hypotheses, we measure pupils' choice of education field at age 25. We classify their career choices via their education field rather than their occupation because many youths at this age are still studying. For each pupil in our sample, we pick the field of the highest degree acquired if they are no longer a student. For anyone who is still studying, we take the field of their current degree. We define fields as primarily female or male dominated based on the generation prior to our sample, i.e. the 13 cohorts who are seven years old during the years 1975-87. If either gender constitutes more than 40% within a field and degree level cell, we define the field as male or female leaning, and gender neutral otherwise.[35] This results in 30% of pupils being in "Male" fields, 43% in gender-neutral, and 27% in "Female" fields. We also report results on STEM and STEM-M (STEM plus Medical) fields as well as teaching-related fields in general and primary school teacher in particular.[36]

Figure 11a and Table 9 report results on the choice of education field. The first three coefficients report results for primarily male, gender neutral and primarily female fields. We can observe a shift away from gender neutral fields towards both more male- and female-dominated fields, but these effects are muted. Turning to STEM and STEM-M, we observe that pupils are more likely to take up such fields when exposed to more male quota teachers, with effect sizes corresponding to a 0.08 and 0.09 SD increase for a 1 SD increase in the share of male quota teachers, respectively.

We then turn to examine results by pupil gender in Figure 11b and Appendix Tables A9 and A10 separately.[37] For both boys and girls, we can observe a shift away from gender neutral fields towards both more male- and female-dominated fields, but only the coefficient of boys moving away from gender neutral fields is significant at a conventional level. In column 4 and 5, we observe coefficients that show a shift towards STEM fields when being exposed to male teachers. Regarding teaching fields overall and primary teacher education specifically, we fail to reject a null effect. The patterns outlined are similar when we control for level of education.

The results for pupils' choice of field are somewhat indicative of both male and female pupils moving away from gender-neutral fields and becoming more inclined to pursue STEM oriented fields when exposed to more male (quota) teachers. As effects move in the same direction for both pupil genders, we again fail to establish support for the hypothesis that effects would be operating through a same-gender role model channel and we do not find evidence for boys – inspired by observing more male teachers during primary school – becoming more likely to take up education related fields or primary school teaching.

7.3 Complementarities between male and female teachers

Based on the predictions outlined in the conceptual framework in section 3, we also try to distinguish whether our results are driven by complementarities between male and female teachers. We do this by assessing marginal returns to more male quota teachers along the distribution of the share of male teachers at baseline (i.e. in 1990): If male and female teachers are comple-

[35] I.e. we define share female based separately for a vocational degree in business vs. an academic degree in business. For the group that has never finished a degree beyond compulsory education and is currently not a student (9.8% of the sample), we assign the gender share of compulsory education, which is categorized as a gender-neutral field based on the previous generation.

[36] We define both academic and vocational degrees as STEM if the three-digit classification of Statistics Finland is one of the following fields: Agriculture (incl. Forestry and Fishery), Biology, Engineering, Environment, ICT, Mathematics and Statistics, Physical Sciences, Veterinary, and the 4-digit category related to Materials (glass, paper, plastic and wood). STEM-M in addition includes the 3-digit field Health.

[37] In all of the regressions on field choices, we do not estimate joint fixed effects for both genders, but report separate regressions by splitting the sample. We do this since for these gendered outcomes, the assumption that shocks would affect boys and girls similarly does not seem justified (i.e. a shock that raises demand for health care workers is likely to have quite different effects on young men vs. young women).

ments, adding an additional male teacher at a place with mostly female teachers should have larger marginal returns compared to adding an additional male teacher in an environment that is close to gender parity. We split the sample by the median of initial share of male teachers in a municipality. The first group has initially a low share of male teachers (on average 29% within the group), and the second group has a relatively higher share of male teachers (on average 43%).[38] Appendix Tables A2 and A3 report results: While the standard errors are large, the estimated coefficients in both groups are similar, and this pattern is consistent across outcomes. We therefore fail to establish strong evidence for the notion that complementarities between male and female teachers are driving our results, but this is mostly due to precision issues. Overall, we cannot convincingly distinguish whether we find positive effects because diversity itself makes teacher teams more productive, or because within the scope of the policy, quota men have higher inherent teaching ability than the marginal women they replace.

8 Robustness

8.1 Do schools change hiring practices due to the reform?

Our treatment coefficients measure the effect of the quota policy, and thus include any impacts that may be due to schools responding endogenously to the policy, for example by changing their hiring patterns and recruiting more experienced teachers in lieu of rookies. While this is not a direct threat to identification, assessing these aspects helps to understand the underlying drivers of our effects. We therefore examine whether any other aspects of local teacher composition changes in response to the reform. Table 10 reports first difference municipal regressions of the form of equation 8 for consistency, with all specifications assessing changes in flows. Our goal is to understand whether retirements that happen in the post-quota period differentially change proxies of teacher experience relative to retirements in the quota period.

In column 3, we examine how retirements affect the average age of the local teacher body: retirements in general reduce the average age among local teachers by 16 years as schools hire younger colleagues to replace retiring teachers. Importantly, we do not detect a sizeable change in the post-quota period: Average age of teachers in response to retirements in the post period drops by an additional 1.3 years, which is not statistically significant and consistent with slightly gendered age profiles when entering the teaching profession observed in section 2.3. In the fourth column, we observe very similar effect sizes for the average number of years since obtaining a teaching degree as a proxy for teaching experience. These results refer to composite changes and could just stem from teacher exit, therefore columns 5 and 6 examine as outcomes two different ways to calculate the share of first entrants to the teaching profession among newly arriving teachers in any given year and municipality.[39] While noisily estimated, retirements in

[38]We can split the sample at other points in the distribution, with results qualitatively similar across different splits. We need to split the sample for this part instead of assessing heterogeneity with a pooled regression since the characteristic by which we assess heterogeneity is constant within municipality, and thus doesn't allow us to instrument properly.

[39]We define as a first entrant any teacher who is first observed in a given year as a teacher and is below the age

general result in a higher share of new entrants among newly arriving teachers at a municipality, but this does not change differentially in the post-quota period. Taken together, we fail to find corroborating evidence for changed re-hiring strategies as a response to the quota reform.

8.2 Heterogeneous treatment effects in twoway fixed effects designs

An active literature has documented that the treatment coefficient of a two-way fixed effects (TWFE) regression, $\hat{\beta}_{fe}$, reports a weighted sum of average treatment effects (ATE) (De Chaisemartin and d'Haultfoeuille, 2020; Arkhangelsky et al., 2021; Athey and Imbens, 2021; Sun and Abraham, 2020; Imai and Kim, 2021; Goodman-Bacon, 2021). If treatment effects are heterogeneous, it is possible that treated observations can receive negative weight such that the sign of the overall effect is of the opposite side than each of the individuals ATEs. Heterogeneous treatment effects in our setting could be an issue if, for example, municipalities that experience a retirement shock early in the post period systematically exhibit a treatment effect of the opposite sign (i.e. an increase in share male), and receive a negative weight.[40] Negative weights by themselves are mechanically the product of any two-way fixed effects specification – it is in combination with heterogeneous treatment effects that problems with sign reversal may arise.

The TWFE literature to date has almost exclusively focused on staggered adoption designs with a binary treatment. An exception are De Chaisemartin and d'Haultfoeuille (2020) who also touch on a more general case where treatment is ordered and finite, such that there are observations that are untreated. De Chaisemartin and d'Haultfoeuille (2020) suggest sensitivity measures that are based on the variance of any weights associated with groups that have a treatment value that is strictly positive. In our setting, and somewhat contrary to the cases tackled in this literature to date, the treatment variable is infinite and zero only for a very few extremely selected observations, which makes the suggested sensitivity measures in De Chaisemartin and d'Haultfoeuille (2020) not directly applicable.

Still, we can follow the reasoning of De Chaisemartin and d'Haultfoeuille, 2020 (also highlighted by Jakiela, 2021), to examine the weights attached to our municipality-by-year groups. Intuitively, groups with lower treatment values should be more likely to receive negative weights as their treatment assignment is below average. The weights derived in De Chaisemartin and d'Haultfoeuille (2020) are proportional to the residual from a regression of the treatment variable on the fixed effects. Assessing the sign of the residuals in our setting is informative about the sign of this weight. We report residuals of regressions for all first stage equations outlined in section 5. For equation 7, this corresponds to:

$$\textit{total share 60 post}_{m,t} = \beta_0 + \beta_1 \textit{total share 60}_{m,t} + \gamma_{m,p} + \eta_t + \epsilon_{m,t} \tag{12}$$

of 28 (column 5) or below the age of 30 (column 6). We examine first entrants as a share of new arrivals since the truncated panel structure of our data results in differential likelihood of observing first entrants and new arrivals at any given point in time, but affects both variables similarly.

[40] Due to for example municipal fixed effects differencing out a larger retirement shock later in the panel such that the earlier event is below the average level of retirements in this municipality.

With *total share 60 post*$_{m,t}$ the share of teachers turning 60 interacted with an indicator for the post period. Appendix Figure A4 plots the mean of binned residuals along the distribution of our treatment variable for equation 7, with colors indicating above (red) or below (blue) median values of treatment intensity. Reassuringly, the weights attached to the municipality-by-year groups follow a clear pattern where observations with below median values of treatment are the ones that on average receive negative weight.

We also probe whether treatment effects in our setting are likely to be heterogeneous across various dimensions. To this extent, we assess whether treatment effects are driven by particular years, municipalities, cohorts or levels of treatment assignment. Appendix Figures A6 and following report leave-one-out specifications, with coefficients not sensitive to dropping particular groups at a time. We also examines whether treatment effects differ across different levels of retirement in the post period (see Appendix Figure A5). While we cannot rule out that there may be some heterogeneity in our estimated treatment effects, given the probing outlined in this section, we think it is unlikely to be sufficient to change the sign of our treatment coefficients.

9 Conclusion

In this paper, we document that a quota that advantaged academically lower scoring men to obtain a study slot for primary teacher education has positive effects on output as measured by their pupils' long run educational and labor market outcomes. We find that pupils who were exposed to a higher share of male quota teachers during their time in primary school are more likely to be either in employment or studying at age 25, and have higher educational attainment as measured by their highest degree achieved.

Using comprehensive register data, we trace these long term effects back to consequential application patterns to post-compulsory education: Pupils exposed to a higher share of male quota teachers are more likely to apply to continue education directly after middle school, to obtain their preferred study slots and to enroll. We show that pupils of either gender apply more in line with their academic abilities, albeit along different margins. We do not find evidence that our main effects are driven by male pupils, ruling out a same-gender role model channel as the main mechanism. We then examine whether more male quota teachers inspire pupils to pursue different educational fields and show that pupils of either gender are more likely to choose a STEM field, but boys in particular are not more likely to choose education or teaching related fields.

The quota in our setting thus did not have an impact along two essential dimensions frequently emphasized in policy debates: It did not contribute to reduce occupational segregation – one of the main drivers of raw gender wage gaps – for the future generation, as boys were not inspired to pick an education-related occupation when exposed to more male teachers via the quota. Neither did the quota contribute to close educational attainment gaps between girls and boys, as both pupil genders benefited equally from exposure to male quota teachers.

Instead, the quota in our setting fixed an inefficiency present in the selection process of teachers. Our results show that the quota succeeded in recruiting male teachers that contributed valuable qualities to the school environment within the parameters of the policy. A promising avenue for future research lies in more thoroughly disentangling whether increased representation affects efficiency via a channel of complementarities in team work, or via channels of inherent candidate ability. A further limitation of our setting is that it explores a partial equilibrium, and we refer to prior (Hsieh et al., 2019; Bleemer, 2021) and future research to gauge the general equilibrium consequences of such policies for the allocation of talent in the wider economy.

Our study directly speaks to concurrent policy issues on affirmative action and optimal selection of candidates and illustrates the importance of carefully considering the relationship between selection criteria and minority status. In settings where a main criterion for choosing candidates is more negatively correlated with a particular group status than with effectiveness on the job, representation targets can help to overcome such misalignment if the selection criteria themselves cannot be easily changed. Both in academia, with current experimentation on SAT requirements for US college applications, and in the private sector, where companies are starting to use balanced candidate lists, such avenues are increasingly being explored. Our results suggest that this may pay off not only in terms of achieving more equitable representation, but also in terms of economic efficiency.

References

Ahern, K. R., & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, *127*(1), 137–197.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Ilr Review*, *30*(2), 175–187.

Antecol, H., Eren, O., & Ozbeklik, S. (2015). The effect of teacher gender on student achievement in primary school. *Journal of Labor Economics*, *33*(1), 63–89.

Arcidiacono, P., & Lovenheim, M. (2016). Affirmative action and the quality-fit trade-off. *Journal of Economic Literature*, *54*(1), 3–51.

Arkhangelsky, D., Imbens, G. W., Lei, L., & Luo, X. (2021). Double-robust two-way-fixed-effects regression for panel data. *arXiv preprint arXiv:2107.13737*.

Athey, S., & Imbens, G. W. (2021). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*.

Baltrunaite, A., Bello, P., Casarico, A., & Profeta, P. (2014). Gender quotas and the quality of politicians. *Journal of Public Economics*, *118*, 62–74.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, *124*(4), 1497–1540.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *science*, *335*(6068), 582–586.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647–713.

Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2018). Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *The Review of Economic Studies*, *86*(1), 191–239.

Bertrand, M., Hanna, R., & Mullainathan, S. (2010). Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics*, *94*(1-2), 16–29.

Besley, T., Folke, O., Persson, T., & Rickne, J. (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden. *American Economic Review*, *107*(8), 2204–42.

Black, S. E., Denning, J. T., & Rothstein, J. (2020). *Winners and losers? The effect of gaining and losing access to selective colleges on education and labor market outcomes* (Working Paper). National Bureau of Economic Research.

Bleemer, Z. (2020). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209.

Bleemer, Z. (2021). *Top percent policies and the return to postsecondary selectivity* (Research & Occasional Paper Series: CSHE. 1.2021). Center for Studies in Higher Education.

Breda, T., Grenet, J., Monnet, M., & Van Effenterre, C. (2018). Can female role models reduce the gender gap in science. *Evidence from classroom interventions in French high schools. HAL, halshs-0173068.*

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, *125*(3), 1101–1144.

Chan, J., & Eyster, E. (2003). Does banning affirmative action lower college student quality? *American Economic Review*, *93*(3), 858–872.

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

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633–79.

Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 1220–1240.

Comi, S., Grasseni, M., Origo, F., & Pagani, L. (2020). Where women make a difference: gender quotas and firms' performance in three European countries. *ILR Review*, *73*(3), 768–793.

Cortés, P., Kasoolu, S., & Pan, C. (2021). *Labor market nationalization policies and exporting firm outcomes: Evidence from Saudi Arabia* (Working Paper). National Bureau of Economic Research.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, *86*(1), 195–210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, *95*(2), 158–165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, *42*(3), 528–554.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

Dhume, S. (2019). India's ethnic quotas are a cautionary tale [2019-01-24]. *The Wall Street Journal*.

Eckbo, B. E., Nygaard, K., & Thorburn, K. S. (2021). Valuation effects of norway's board gender-quota law revisited. *Management Science*.

Ellison, G., & Pathak, P. A. (2021). The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from Chicago's exam schools. *American Economic Review*, *111*(3), 943–75.

Etelä Suomen Sanomat. (1988). Koulutuspaikkojen lisäzksellä eroon opettajapulasta [1988-08-25]. *Etelä Suomen Sanomat*.

Eurostat. (2021). Education and training in the EU - facts and figures [Online; accessed 21-Sep-2021].

Ferrari, G., Ferraro, V., Profeta, P., & Pronzato, C. (2021). Do board gender quotas matter? Selection, performance and stock market effects. *Management Science*.

Finnish National Agency for Education. (2018). *Finnish teachers and principals in figures*. Finnish National Agency for Education.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Hanushek, E. A., & Rivkin, S. G. (2006). Teacher quality. *Handbook of the Economics of Education*, *2*, 1051–1078.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and US economic growth. *Econometrica*, *87*(5), 1439–1474.

Imai, K., & Kim, I. S. (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, *29*(3), 405–415.

Izadi, R. (2021). *Teacher selection in finland* (Unpublished manuscript).

Jakiela, P. (2021). Simple diagnostics for two-way fixed effects. *arXiv preprint arXiv:2103.13229*.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? evidence from New York City. *Economics of Education review*, *27*(6), 615–631.

Kivinen, O., & Rinne, R. (1994). The thirst for learning, or protecting one's niche? the shaping of teacher training in Finland during the 19th and 20th centuries. *British journal of sociology of education*, *15*(4), 515–527.

Kofoed, M. S. et al. (2019). The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at West Point. *Journal of Human Resources*, *54*(2), 430–467.

Leman, N. (2021). Can affirmative action survive? [2021-07-26]. *The New Yorker*.

Liiten, M. (2012). Opetus- ja sosiaalialan koulutukseen harkitaan mieskiintiöitä [2012-10-06]. *Helsingin Sanomat*.

Lim, J., & Meer, J. (2017). The impact of teacher–student gender matches random assignment evidence from South Korea. *Journal of Human Resources*, *52*(4), 979–997.

Long, H. (2019). 80 nations set quotas for female leaders. Should the U.S. be next? [2019-03-29]. *Washington Post*.

Lundberg, S. J., & Startz, R. (1983). Private discrimination and social intervention in competitive labor market. *The American Economic Review*, *73*(3), 340–347.

Malinen, O.-P., Väisänen, P., & Savolainen, H. (2012). Teacher education in finland: A review of a national effort for preparing teachers for the future. *Curriculum Journal*, *23*(4), 567–584.

Mankki, V., Mäkinen, M., & Räihä, P. (2020). Teacher educators' predictability and student selection paradigms in entrance examinations for the Finnish primary school teacher education programme. *European Journal of Teacher Education*, *43*(2), 151–164.

Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, *5*(3), 136–69.

Niemi, H., Toom, A., & Kallioniemi, A. (2016). *Miracle of education: The principles and practices of teaching and learning in Finnish schools*. Springer.

OECD. (2019). Society at a glance 2019: Oecd social indicators [[Online; accessed 21-Sep-2021]].

OECD. (2021). Oecd employment outlook [[Online; accessed 21-Sep-2012]].

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, *62*(4), 659–661.

Porter, C., & Serra, D. (2019). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*.

Räihä, P. (2010). Koskaan et muuttua saa! Luokanopettajakoulutuksen opiskelijavalintojen uudistamisen vaikeudesta. *Tampere University Press*.

Riise, J., Willage, B., & Willén, A. (2020). Can female doctors cure the gender STEMM gap? evidence from exogenously-assigned general practitioners. *The Review of Economics and Statistics*, 1–45.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Sahlberg, P. (2021). *Finnish lessons 3. 0: What can the world learn from educational change in Finland?* Teachers College Press.

Sahlberg, P., Johnson, P., & Strauss, V. (2019). What Finland is really doing to improve its acclaimed schools. *Washington Post*.

Saloviita, T. (2018). How common are inclusive educational practices among finnish teachers? *International Journal of Inclusive Education*, *22*(5), 560–575.

Sansone, D. (2017). Why does teacher gender matter? *Economics of Education Review*, *61*, 9–18.

Siniscalchi, M., & Veronesi, P. (2021). *Self-image bias and lost talent* (Working Paper). National Bureau of Economic Research.

Sun, L., & Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

Sysiharju, A.-L. (1987). Women school employees in Finland. In P. A. Schmuck (Ed.), *Women educators: Employees of schools in Western world countries* (pp. 21–41). Suny Press.

Tasa-Arvovaltuutetu. (1987). Tasa-Arvovaltuutetun toimisto 126/1987, lausunto kiintiöiden käytöstä [Office of the High Commissioner of Equality (Paavo Nikula): Opinion on quotas 126/1987, obtained by private correspondence with the Office of the Equality Ombudsman].

Tirri, K. (2014). The last 40 years in Finnish teacher education. *Journal of Education for Teaching*, *40*(5), 600–609.

UN. (2019). Progress on the sustainable development goals: The gender snapshot 2019.

Uusiautti, S., & Määttä, K. (2013). Significant trends in the development of Finnish teacher education programs (1860-2010).

Uusiautti, S., Määttä, K. et al. (2013). Who is a suitable teacher? the over-100-year long history of student selection for teacher training in Finland. *International Journal of Sciences*, *2*(03), 109–118.

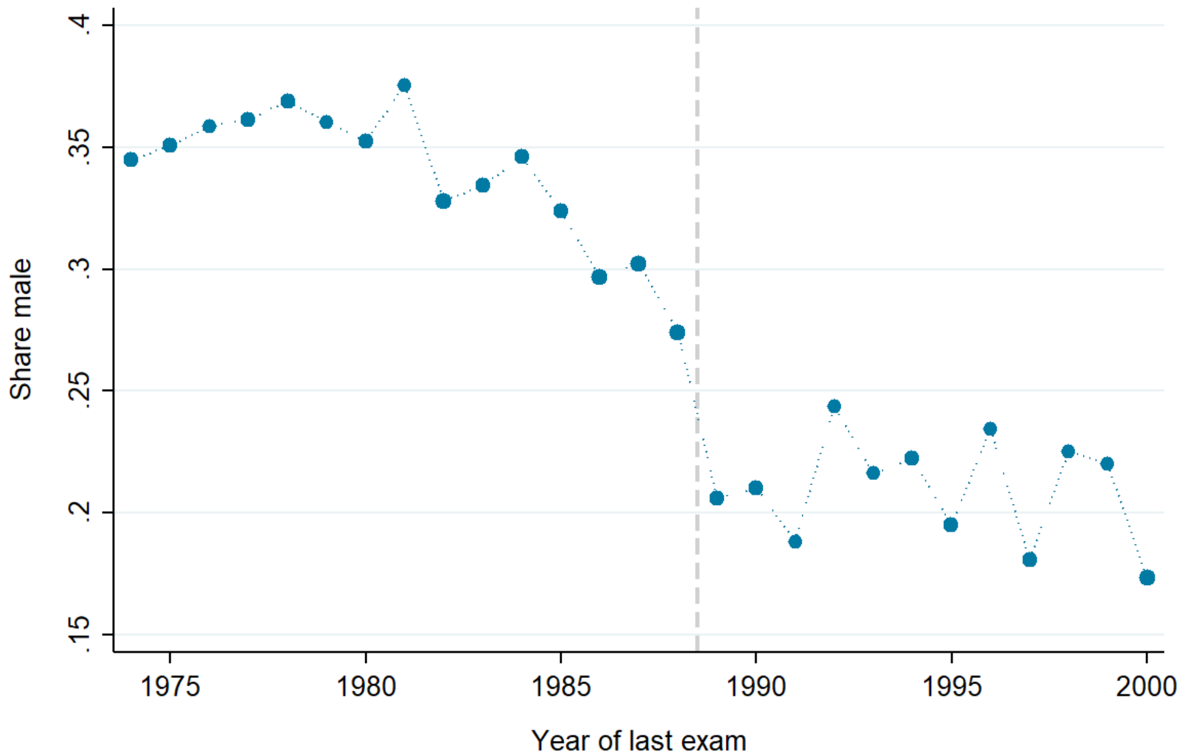
Virtanen, H. (2016). Essays on Post-Compulsory Education Attainment in Finland.

Welch, F. (1976). Employment quotas for minorities. *Journal of Political Economy*, *84*(4, Part 2), S105–S141.

Yagan, D. (2016). Supply vs. demand under an affirmative action ban: Estimates from UC law schools. *Journal of Public Economics*, *137*, 38–50.

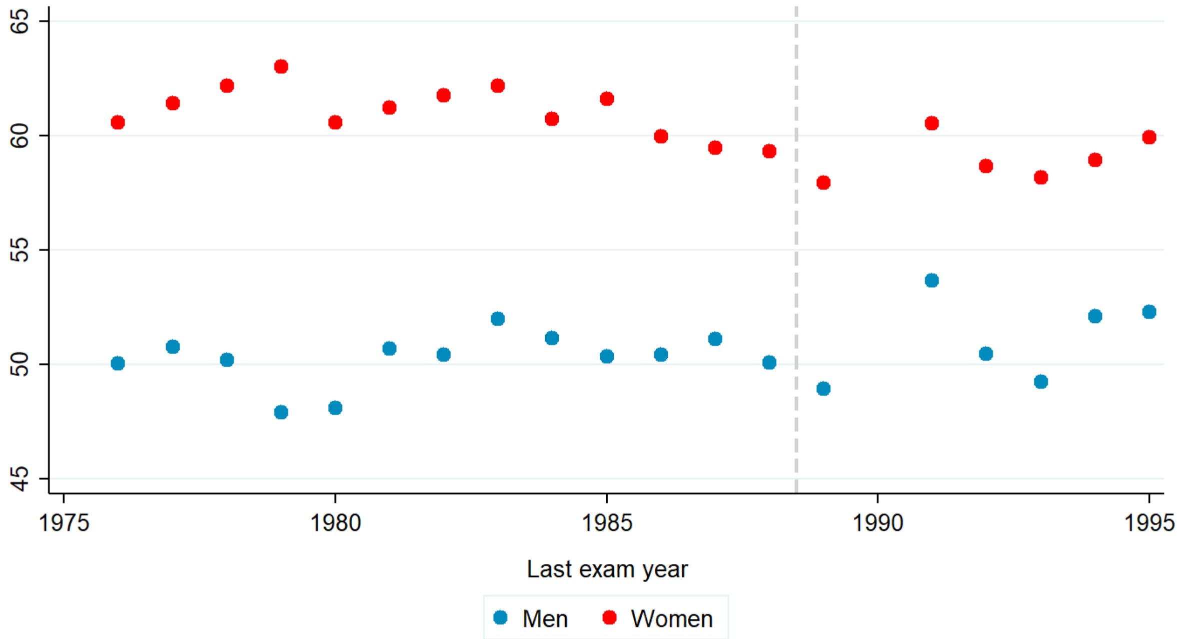
Tables and Figures

Figure 1: Share male among primary school teaching degree holders by year of last matriculation exam



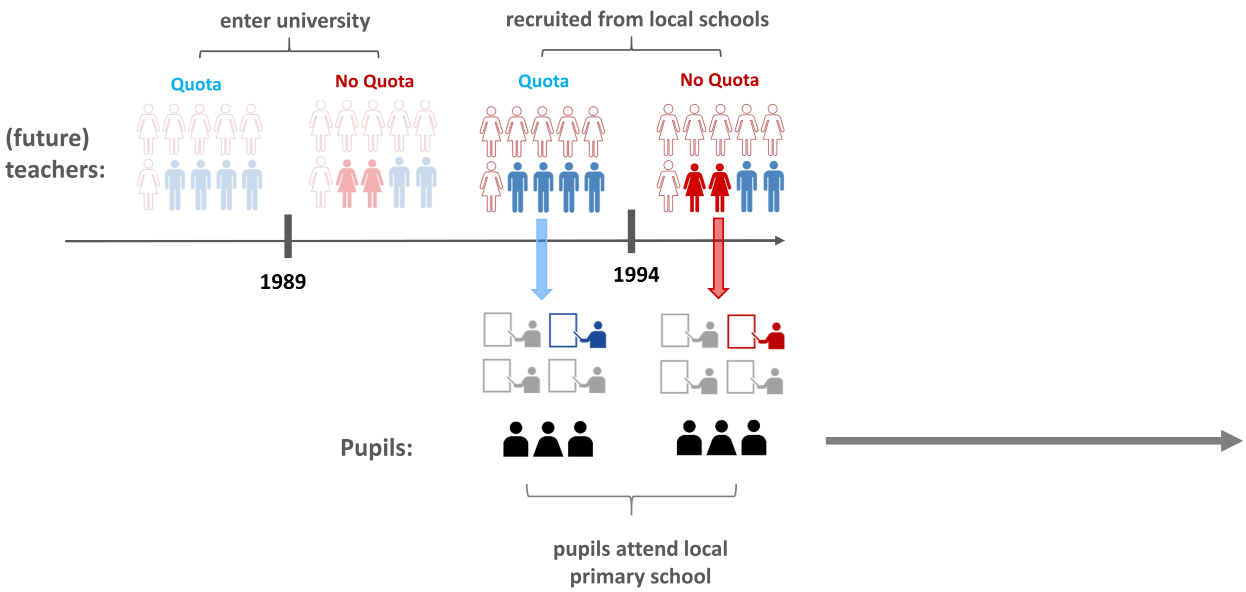
Note: Share male among primary school teaching degree holders, by the last year in which they took the matriculation exam, which qualifies applicants for university admissions. Exam takers in 1989 (dashed grey line) and thereafter will have studied after the male quota was abolished. (back)

Figure 2: Matriculation exam percentile rank among primary school teaching degree holders by year of last matriculation exam



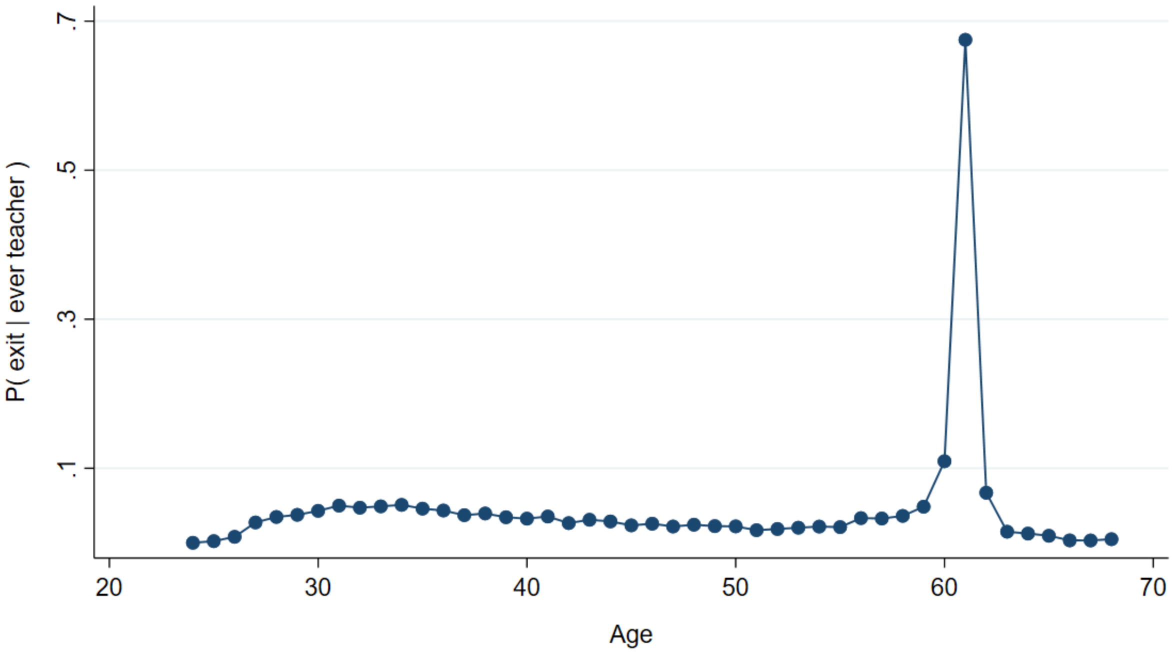
Note: National percentile rank across all subjects in the matriculation exam among primary school teaching degree holders, by gender and the last year in which they took the matriculation exam (qualifies applicants for university admissions). Exam takers in 1989 (dashed grey line) and thereafter will have studied after the male quota was abolished. Data on exam points for the year 1990 are missing, so that we cannot calculate the national distribution of scores according to percentiles for that year. Note that if the worst scorers are those that repeat the exam, this will bias the average scores in the years after 1990 upwards (the worst performers will not be counted in these averages as their scores are missing). We address this issue by also examining average grades, for which we have data in 1990. We can show that the pattern is very similar and upward bias in the percentile scores after 1990 should be small. (back)

Figure 3: Timeline of the reform



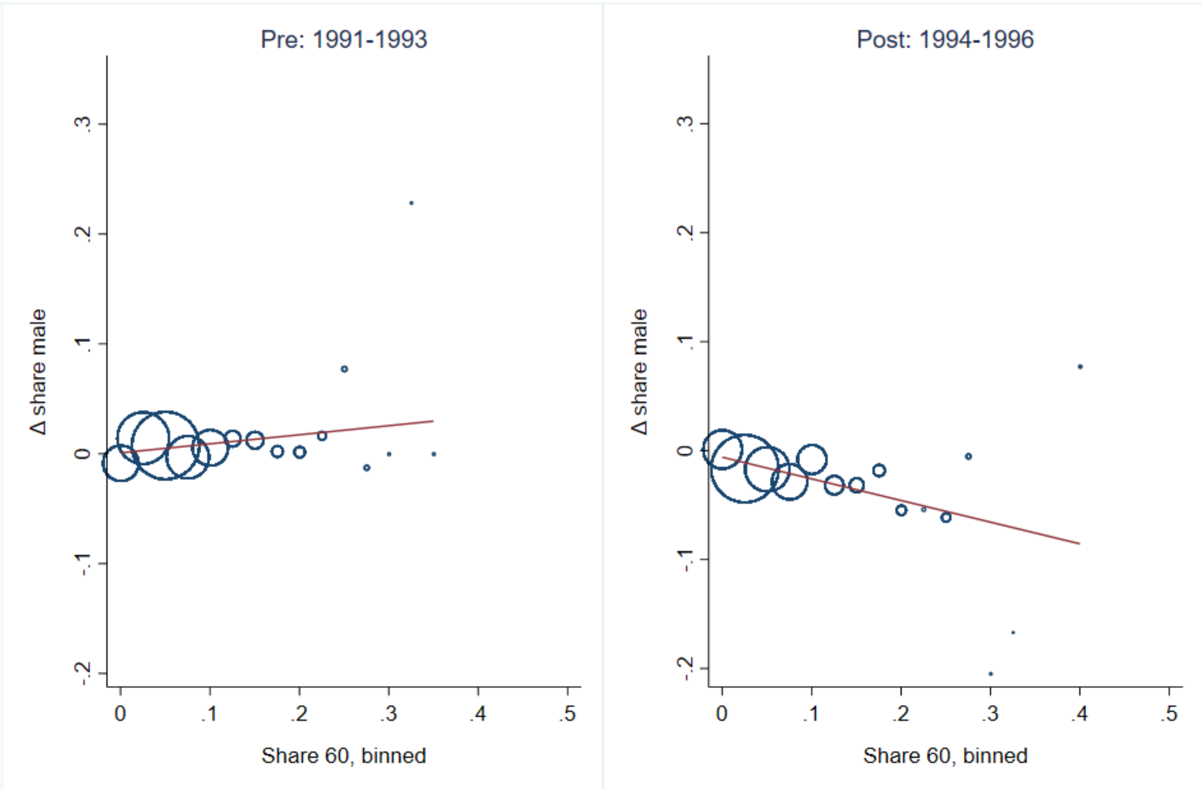
Note: Future primary school teachers enter university with the quota (pre-1989) and without the quota (1989 and thereafter), and graduate from the five year primary school teaching degree around the year 1994. Primary teacher graduates get hired by municipalities to teach in local schools. Pupils will experience differential exposure to quota teachers, described in detail in section 5 of the paper. (back)

Figure 4: Probability of teacher exit by age



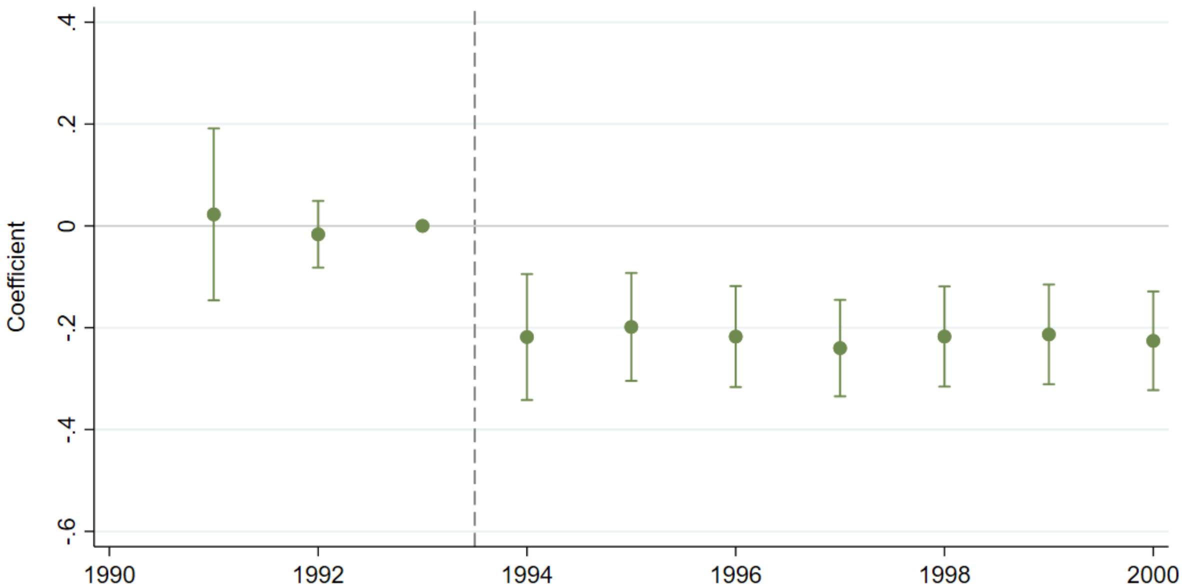
Note: Share of primary school teachers not working as a primary school teacher at a given age, conditional on having worked as a primary school teacher in the previous year. Data for all active primary school teachers in the years 1990-2000. Multiple exits per teacher possible.
(back)

Figure 5: First stage intuition: Changes in share male teachers by local retirements, raw data



Note: Municipality level data, binned: Change in the share of male primary school teachers for a period of similar length in the quota (1991-93) and post-quota (1994-96) period against total share of teachers turning 60. Linear fit, weighted by the number of municipalities per bin.
(back)

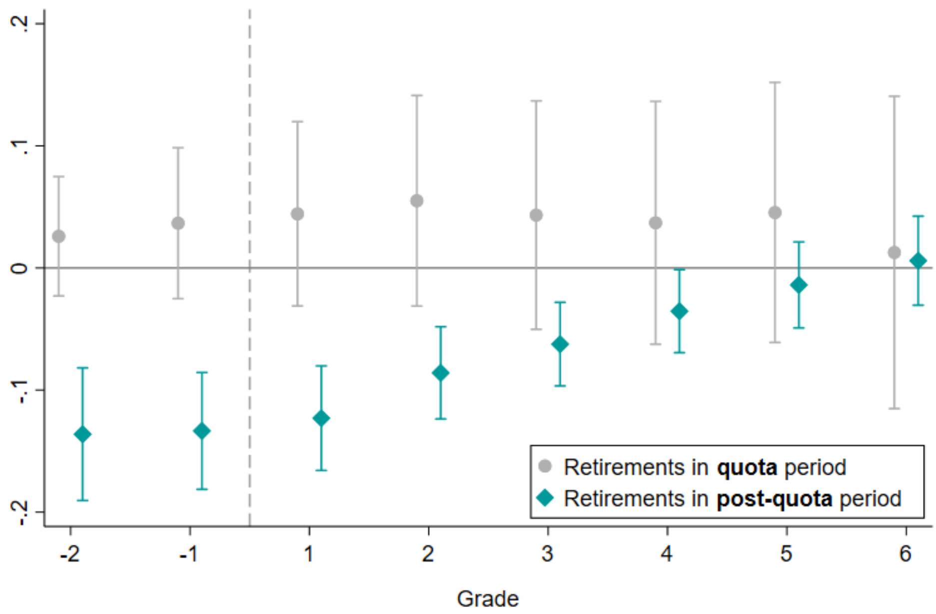
Figure 6: Municipal level first stage as event study



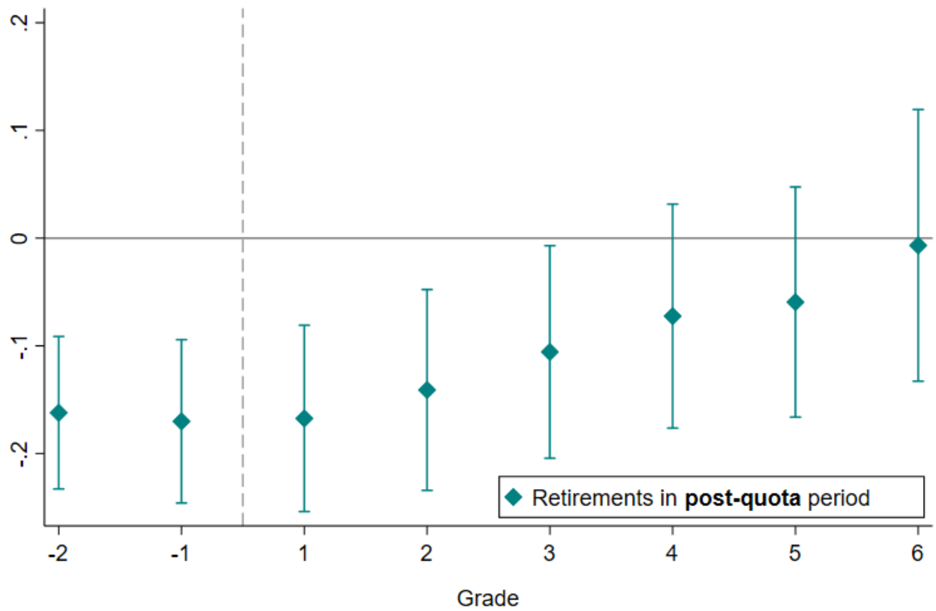
Note: Year-on-year estimates of π_2 for the fixed-effects first stage Equation 7. Standard errors clustered at the municipality level. Population weighted. (back)

Figure 7: First stage pupil level

(a) Separate estimation of quota and post-quota coefficients



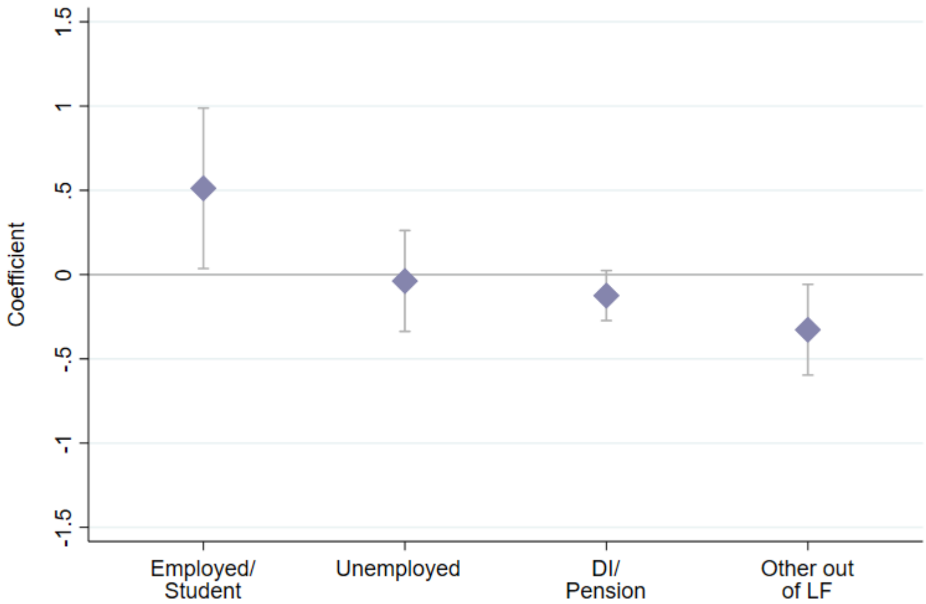
(b) Post-quota (relative to quota coefficients)



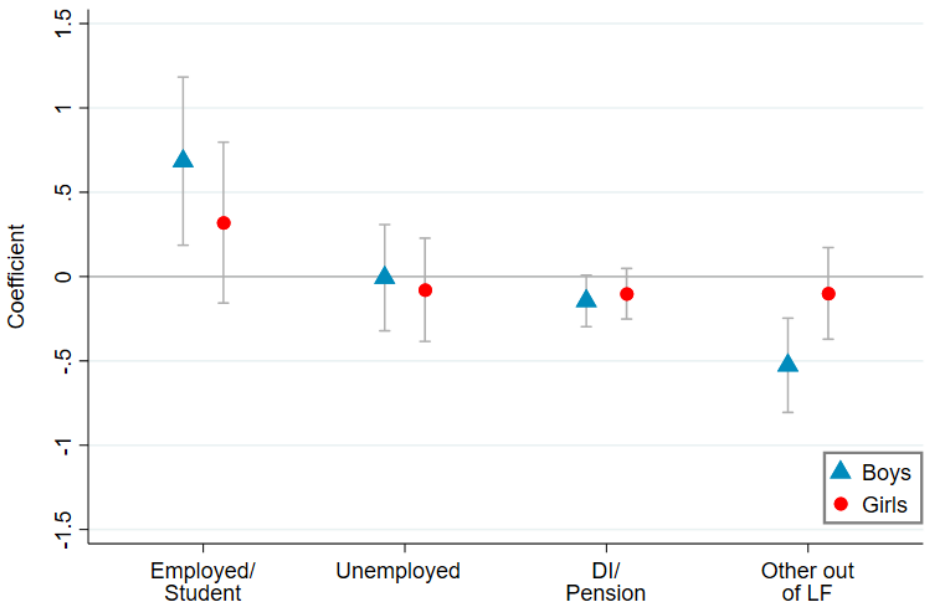
Note: Grade level estimation of pupil level first stage (c.f. Equation 11). Panel (a) estimates absolute coefficients for effect of retirements pupils experience in the quota and the post-quota period. Panel (b) depicts coefficients for the post-quota period *relative* to the quota period (i.e. it shows the difference between quota and post-quota estimates depicted in Panel (a)). Outcome is the average share of male teachers a pupil is exposed to during their time in primary school (Grades 1-6). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. (back)

Figure 8

(a) Labor market outcomes at age 25



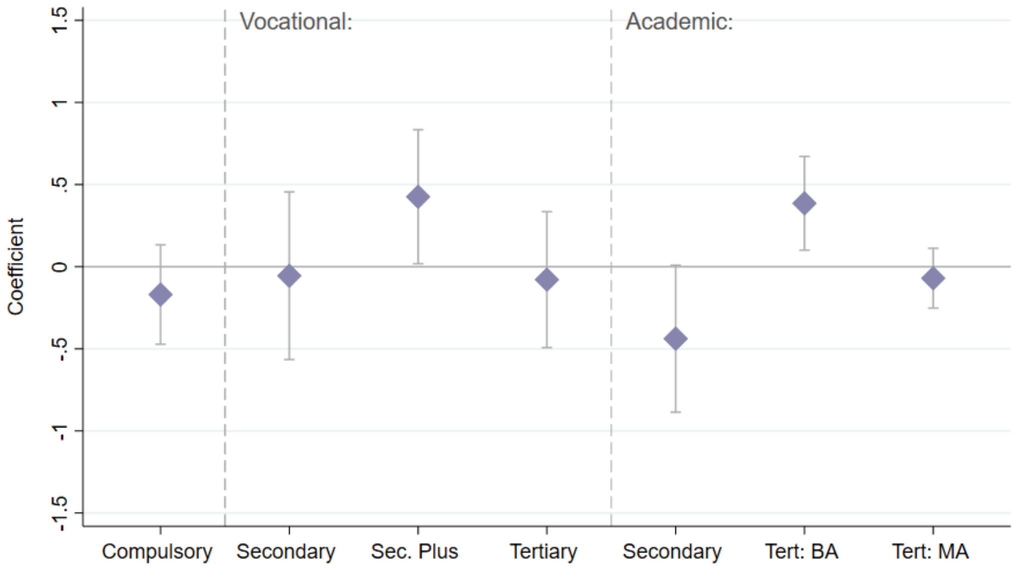
(b) By gender: Labor market outcomes at age 25



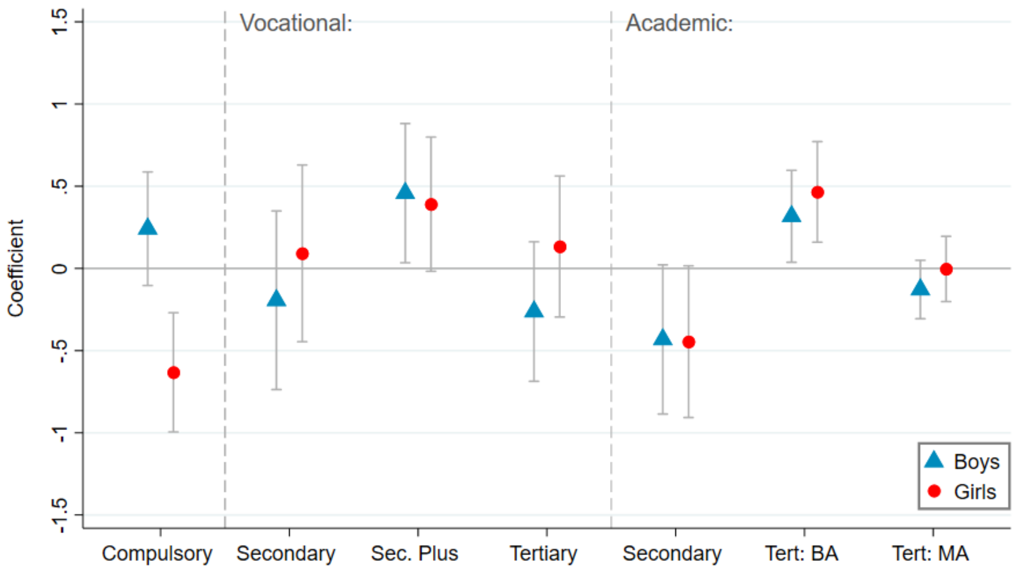
Note: IV estimates of Equation 9. Outcome: Mutually exclusive categories of labor market status in last week of the year in which pupils turn 25 years old. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.
(back)

Figure 9

(a) Highest degree achieved at age 25



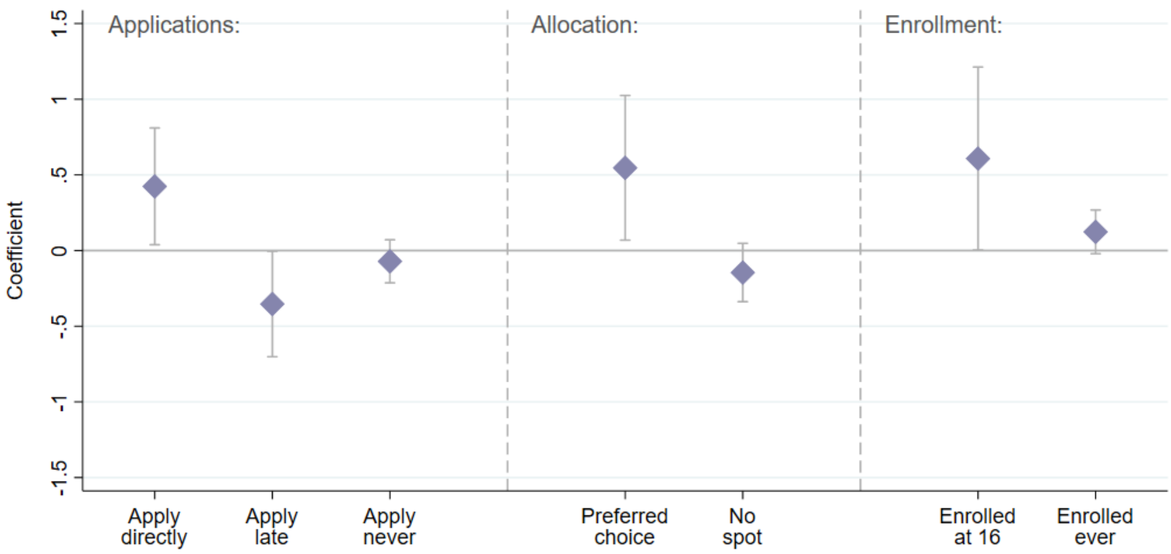
(b) By gender: Highest degree achieved at age 25



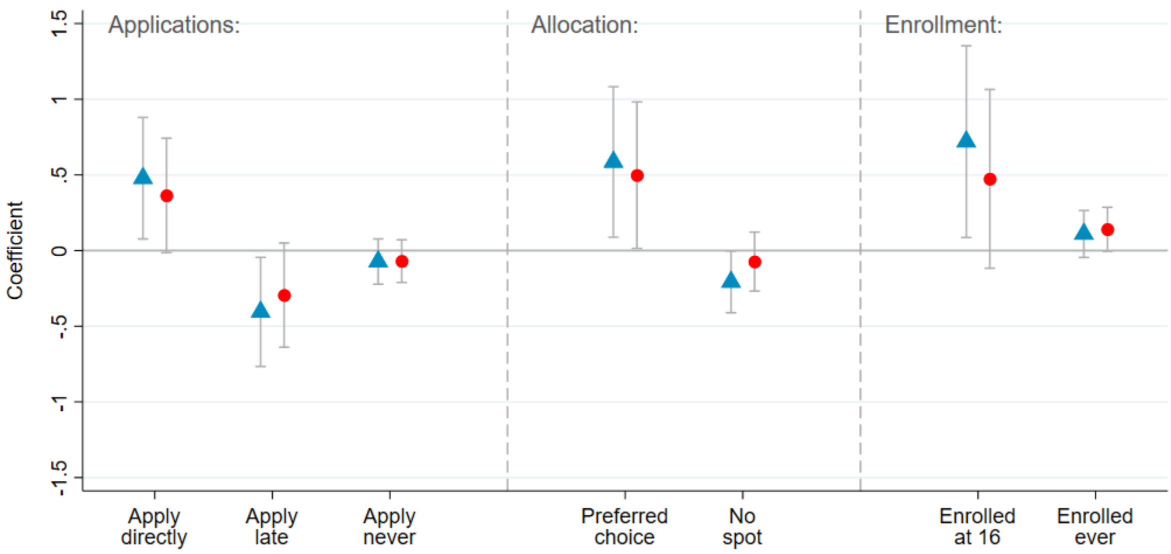
Note: IV estimates of Equation 9. Outcomes are mutually exclusive categories of pupils' highest degree achieved at age 25: Having only Compulsory education. For the Vocational track: Having a basic three year secondary degree (Secondary), having additional qualifications or high school coursework beyond a basic degree (Sec. Plus), having a tertiary degree from a polytechnic (Tertiary). For the Academic track: Having a three year high school degree (Secondary), having a three year university BA degree (Tert: BA), having a two year university MA degree (Tert: MA) or higher. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. (back)

Figure 10

(a) Applications and enrollment for upper secondary education



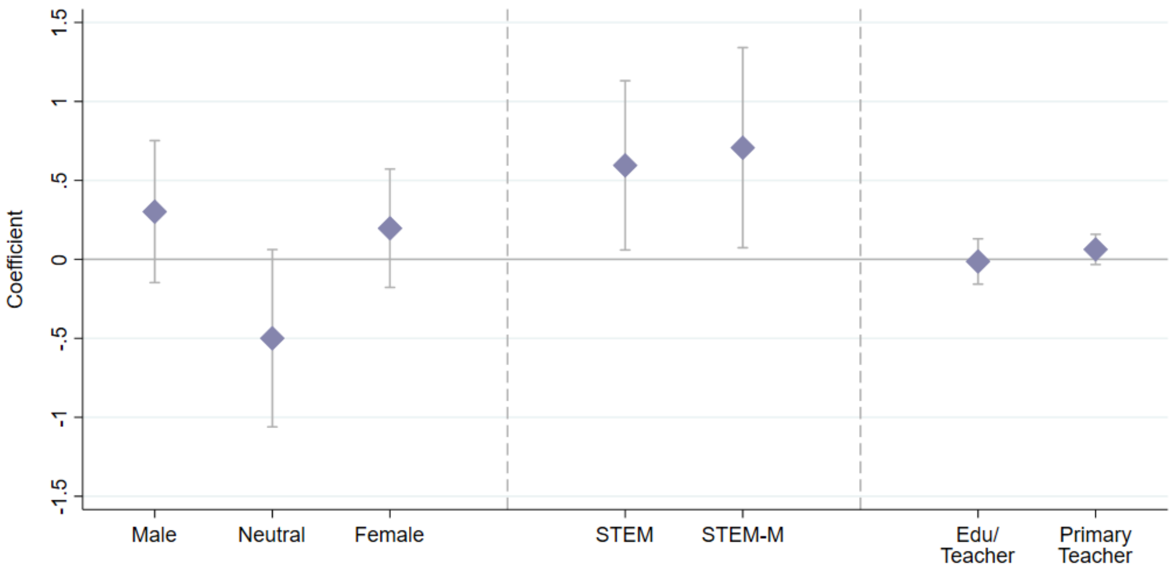
(b) By gender: Applications and enrollment for upper secondary education



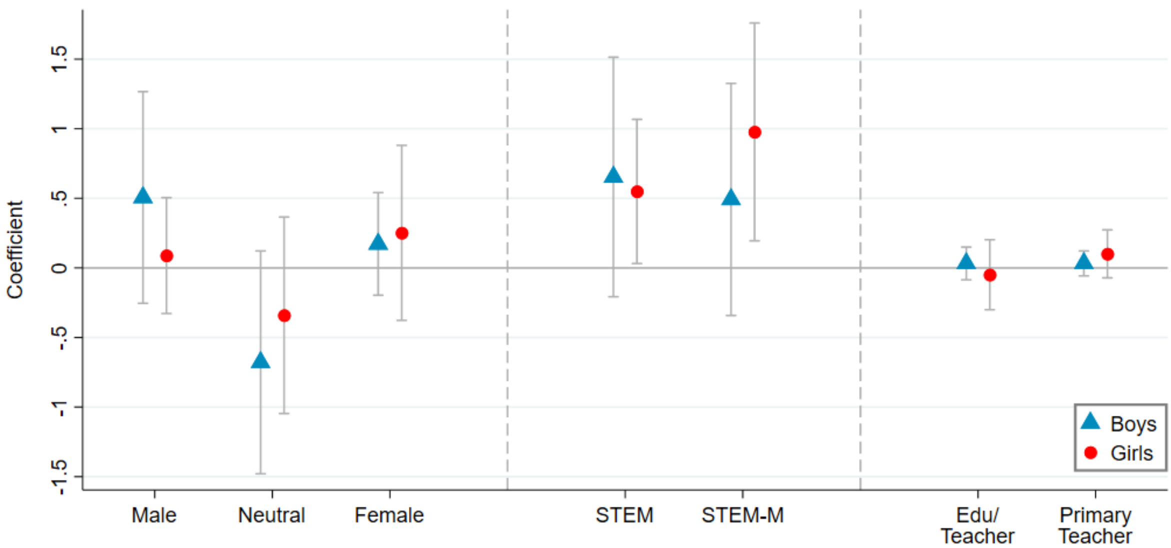
Note: IV estimates of Equation 9. Outcomes from left to right: "Applications" to upper secondary education (mutually exclusive categories): Pupils apply directly in spring of the year in which they turn 16 (Apply directly), they apply up to four years after they have turned 16 (Apply late), or they apply never or later than five years after having turned 16 (Apply never). "Allocation": Pupils obtain one of their first two preferred choices in the application (Preferred choice), or do not obtain a study slot (No spot). "Enrollment": Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at age 16), and ever enrolled in upper secondary education up to age 25 (Ever enrolled). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. (back)

Figure 11

(a) Education field at age 25



(b) By gender: Education field at age 25



Note: IV estimates of Equation 9, separate regressions by gender. Outcomes from left to right: Field is 'Male' dominated ($\geq$ 40% male), (gender) 'Neutral' or 'Female' dominated ($\geq$ 40% female). Field is STEM or STEM + Medicine (STEM-M). Field is Education Science or Teacher. Field is Primary School Teacher. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. (back)

Table 1: Summary statistics of teachers by gender

	Quota			Post Quota		
Variable	Female	Male	Difference	Female	Male	Difference
Background and current place of living						
Urban residence at birth	0.586	0.562	-0.023***	0.612	0.590	-0.022***
			(0.007)			(0.007)
Rural residence at birth	0.371	0.392	0.021**	0.358	0.376	0.018**
			(0.007)			(0.006)
Born on Russian territory	0.043	0.046	0.003	0.028	0.033	0.005*
			(0.003)			(0.002)
Finnish mother tongue	0.922	0.947	0.025***	0.916	0.940	0.024***
			(0.004)			(0.004)
Lives in region of birth	0.457	0.488	0.031***	0.473	0.496	0.023***
			(0.007)			(0.007)
Lives in municipality of birth	0.229	0.268	0.039***	0.249	0.277	0.028***
			(0.006)			(0.006)
Education path (born after 1952)						
1.05						
High school degree	0.980	0.979	-0.002	0.983	0.980	-0.003
			(0.003)			(0.003)
Teaching degree	0.894	0.889	-0.005	0.920	0.910	-0.010
			(0.007)			(0.006)
Age at high school degree	19.25	19.47	0.23***	19.23	19.46	0.23***
			(0.02)			(0.02)
Age at teacher degree	25.54	26.50	0.96***	25.48	26.48	1.00***
			(0.09)			(0.09)
Academic performance (born after 1952)						
Matriculation exam	0.984	0.981	-0.003	0.935	0.948	0.013**
			(0.003)			(0.004)
National percentile rank, first take	62.85	50.73	-12.12***	61.01	50.46	-10.54***
			(0.36)			(0.31)
National percentile rank, best take	63.86	52.40	-11.45***	62.00	52.05	-9.96***
			(0.34)			(0.30)
Mathematics exam	0.741	0.826	0.086***	0.768	0.842	0.074***
			(0.009)			(0.007)
Advanced mathematics exam	0.277	0.380	0.103***	0.263	0.380	0.117***
			(0.010)			(0.008)

Note: Characteristics of male and female primary school teachers who are active teachers for at least one year in the quota period (1990-93) or in the post quota period (1994-2000) and who are between 24 and 60 years old. (back)

Table 2: First stage at the municipal level

	First Differences				Fixed Effects		
Dependent variable	Δ male				share male		
Share 60	0.062	0.062	0.074*	0.072*			
	(0.038)	(0.039)	(0.041)	(0.039)			
Share 60 * post	–0.165***	–0.170***	–0.175***	–0.161***			
	(0.044)	(0.044)	(0.047)	(0.044)			
1.05 Total share 60					0.068	0.086*	0.078*
					(0.043)	(0.045)	(0.043)
Total share 60 * post					–0.218***	–0.232***	–0.194***
					(0.049)	(0.054)	(0.049)
M*Post FE					X	X	X
Y FE		X		X	X		X
R*Y FE			X			X	
M*Y controls				X			X
Adj. R^2	0.017	0.022	0.011	0.025	0.869	0.866	0.869
Obs	4448	4448	4439	4448	4443	4434	4443
Dep mean	.0007	.0007	.0007	.0007	.3601	.3601	.3601

Note: Estimates for Equation 8 (columns 1-4) and Equation 7 (columns 5-7). Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted. Time-varying municipal controls include log population, log household income, share unemployed, share of adult population with compulsory, secondary and tertiary education, share of families in single parent HH. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 3: First stage: Teachers' matriculation exam scores

	Total	Language	Math	Science
Total share 60	–1.07	–4.27	–3.51	–1.09
	(4.25)	(5.09)	(7.67)	(5.45)
Total share 60 * post	1.25	3.02	1.28	1.95
1.05	(4.70)	(5.68)	(8.55)	(6.08)
M*Post FE	X	X	X	X
Y FE	X	X	X	X
Adj. R^2	0.85	0.86	0.80	0.82
Obs	4329	4329	4314	4317
Dep mean	58.88	60.92	52.17	61.93

Note: Estimates for Equation 7 with average of local teachers' national percentile rank in first attempt of matriculation exam as the outcome. Language includes scores for mother tongue (FI/SE) and second national language (SE/FI). Science (Reaali) scores include the combined scores across subjects ranging from history and religion to chemistry and physics. Data available only for teacher cohorts born after 1952. Sample is restricted to municipalities where there is at least one teacher with observed score in 1991. Standard errors clustered at the municipality level. Regressions weighted by population. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 4: Pupil level first stage, reduced form and IV for being Employed/Student at age 25

	First stage			RF			IV		
	Avg share male			Employed or student			Employed or student		
Total share 60	0.034	0.034	0.044	0.031	0.032	0.042	0.021	0.019	0.020
	(0.036)	(0.036)	(0.037)	(0.028)	(0.028)	(0.029)	(0.026)	(0.027)	(0.027)
Total share 60 *	−0.177***	−0.177***	−0.168***	−0.053	−0.071**	−0.086***			
1.05 post	(0.042)	(0.042)	(0.043)	(0.033)	(0.032)	(0.033)			
Avg share male							0.297	0.403*	0.512**
							(0.215)	(0.216)	(0.243)
Municipal FE	X	X	X	X	X	X	X	X	X
Cohort FE	X	X		X	X		X	X	
Region*Cohort FE			X			X			X
Ind. controls		X	X		X	X		X	X
MOP F^{eff}							17.96	17.94	15.37
Adj. R^2	0.916	0.916	0.921	0.008	0.025	0.025			
Obs	811,392	811,392	811,392	811,392	811,392	811,392	811,392	811,392	811,392
Dep mean	.3133	.3133	.3133	.8423	.8423	.8423	.8423	.8423	.8423

Note: Columns 1-3 show estimates for Equation 11 with the average share male teachers pupils are exposed to during primary school as the outcome. Columns 4-6 show reduced form estimates (corresponding to 11), and Columns 7-9 show IV estimates of Equation 9 with being either employed or a student at age 25 as the outcome. Standard errors clustered at the municipality level. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 5: Labor market outcomes at age 25

1.05

	Employed/ Student	Un- employed	DI/ Pension	Other out of LF
Avg share male	0.512**	–0.038	–0.124	–0.327**
	(0.243)	(0.153)	(0.076)	(0.137)
Municipal FE	X	X	X	X
Region*Cohort FE	X	X	X	X
Ind. controls	X	X	X	X
MOP F^{eff}	15.37	15.37	15.37	15.37
Obs	811,392	811,392	811,392	811,392
Dep mean	.842	.086	.017	.053
Std effect	.089	-.009	-.061	-.093

Note: Estimates for Equation 9. Outcomes are mutually exclusive categories of pupils' labor market status measured at age 25: Being in employment or a student, unemployed, on disability insurance (DI) or receiving pension payments, or being out of the labor force for other reasons. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 6: Highest degree achieved at age 25

1.05 1.05

	Compulsory	Vocational			Academic		
	schooling	Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Avg share male	–0.169	–0.055	0.426**	–0.079	–0.438*	0.386***	–0.070
	(0.154)	(0.260)	(0.208)	(0.211)	(0.228)	(0.146)	(0.093)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.36	15.36	15.36	15.36	15.36	15.36	15.36
Obs	810,065	810,065	810,065	810,065	810,065	810,065	810,065
Dep mean	.127	.316	.108	.146	.211	.054	.038
Std effect	-.032	-.008	.088	-.014	-.068	.108	-.023

Note: IV estimates of Equation 9. Outcomes are mutually exclusive categories of pupils' highest degree achieved at age 25: Having only Compulsory education. For the Vocational track: Having a basic three year secondary degree (Sec), having additional qualifications or high school coursework beyond a basic degree (Sec Plus), having a tertiary degree from a polytechnic (Tert). For the Academic track: Having a three year high school degree (Sec), having a three year university BA degree (Tert: BA), having a two year university MA degree (Tert: MA) or higher. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 7: First stage, reduced form and IV: Applications in last year of middle school

	First stage			RF			IV		
	Avg share male			Apply			Apply		
Total share 60	0.033	0.033	0.043	0.048**	0.048**	0.018	0.031	0.025	–0.001
	(0.036)	(0.036)	(0.037)	(0.020)	(0.021)	(0.021)	(0.025)	(0.027)	(0.020)
Total share 60 * post	–0.176***	–0.176***	–0.168***	–0.096***	–0.122***	–0.071***			
1.05	(0.042)	(0.042)	(0.043)	(0.026)	(0.031)	(0.027)			
Avg share male							0.544***	0.696***	0.424**
							(0.202)	(0.234)	(0.197)
Municipal FE	X	X	X	X	X	X	X	X	X
Cohort FE	X	X		X	X		X	X	
Region*Cohort FE			X			X			X
Ind. controls		X	X		X	X		X	X
MOP F^{eff}							17.66	17.64	15.28
Adj. R^2	0.916	0.916	0.922	0.038	0.069	0.070			
Obs	825,094	825,094	825,094	825,094	825,094	825,094	825,094	825,094	825,094
Dep mean	.313	.313	.313	.9106	.9106	.9106	.9106	.9106	.9106

Note: Columns 1-3 show estimates for Equation 11 with the average share male teachers pupils are exposed to during primary school as the outcome. Columns 4-6 show reduced form estimates (corresponding to Equation 11), and Columns 7-9 show IV estimates of Equation 9, with a pupil applying directly in the spring of the year they turn 16 (i.e. the last year of middle school) as the outcome. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 8: Applications and enrollment for post-compulsory education

	Apply directly	Apply late	Apply never	Pref. choice	No spot	Enrolled at 16	Enrolled ever
Avg share male	0.424**	−0.353**	−0.071	0.547**	−0.145	0.608**	0.124*
	(0.197)	(0.178)	(0.073)	(0.244)	(0.098)	(0.309)	(0.074)
1.05 Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.28	15.28	15.28	15.28	15.28	13.00	13.00
Obs	825,094	825,094	825,094	825,094	825,094	695,340	695,340
Dep mean	.911	.067	.022	.862	.04	.861	.98
Std effect	.095	-.09	-.031	.101	-.047	.11	.055

Note: IV estimates for Equation 9. Outcomes in columns 1-3 are mutually exclusive categories of applications to upper secondary education: Pupils apply directly in spring of the year in which they turn 16 (Apply directly), they apply up to four years after they have turned 16 (Apply late), or they apply never or later than five years after having turned 16 (Apply never). "Allocation" (columns 4-5): Pupils obtain one of their first two preferred choices in the application (Pref. choice), or do not obtain a study slot (No spot). "Enrollment" (columns 6-7): Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at age 16), and ever enrolled in upper secondary education up to age 25 (Ever enrolled). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 9: All: Field of education at age 25

1.05

	Male	Neutral	Female	STEM	STEM-M	Education/ Teacher	Primary Teacher
Avg share male	0.302	–0.500*	0.197	0.595**	0.707**	–0.013	0.063
	(0.229)	(0.286)	(0.191)	(0.273)	(0.323)	(0.073)	(0.049)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.37	15.37	15.37	15.37	15.37	15.37	15.37
Obs	811,392	811,392	811,392	811,392	811,392	811,392	811,392
Dep mean	.303	.433	.264	.264	.379	.023	.011
Std effect	.042	-.064	.028	.086	.093	-.006	.039

Note: IV estimates of Equation 9. Outcomes from left to right: Field is 'Male' dominated ($\geq$ 40% male), (gender) 'Neutral' or 'Female' dominated ($\geq$ 40% female), based on previous generation. Field is STEM or STEM + Medicine (STEM-M). Field is Education Science or Teacher. Field is Primary School Teacher. Individual level controls are measured at age 7 and include language (SE/FI), foreign origin, single parent HH, highest degree attained in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table 10: Exit and hiring patterns in municipalities

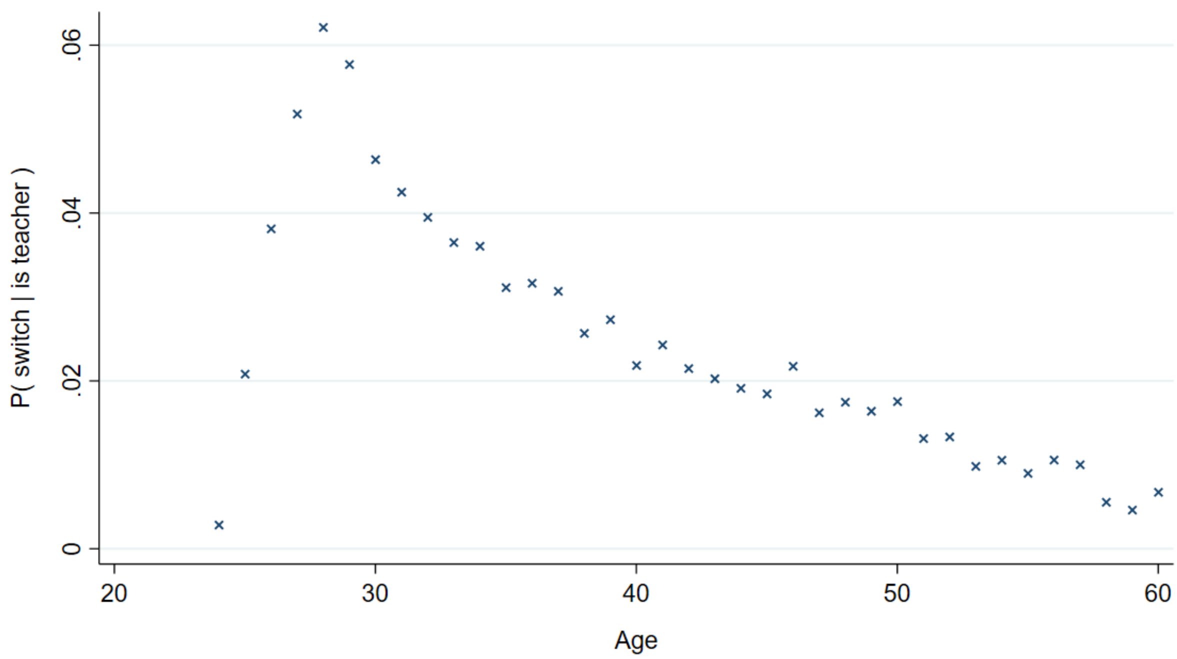
1.05

	Leave	Leave ≤ 55	Δ Age	Δ Time since degree	First entr (≤ 28)	First entr (≤ 30)
Share 60	0.873***	–0.049	–16.088***	–15.599***	0.335	0.386
	(0.061)	(0.057)	(1.215)	(1.154)	(0.315)	(0.314)
Share 60 * post	0.041	–0.000	–1.300	–1.051	–0.043	–0.126
	(0.067)	(0.063)	(1.387)	(1.323)	(0.375)	(0.372)
Y FE	X	X	X	X	X	X
Adj. R^2	0.176	0.011	0.222	0.211	0.038	0.036
Obs	4448	4448	4448	4448	3746	3746
Dep mean	.1	.07	-.21	-.2	.35	.4

Estimates for Equation 8. Outcomes from left to right are: Share of teachers exiting, share of teachers below age 55 exiting, year-on-year changes in average age of all local teachers (Δ Age), average time since obtaining a teaching degree of all local teachers (Δ Time since degree). The share of new teacher arriving that are first entrants defined as not having taught before and being below age 28 (column 5), and below age 30 (column 6). Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted. * p<0.1, ** p<0.05, *** p<0.01 (back)

Appendix

Figure A1: Probability of switching municipality of work for active teachers by age



Note: Share of primary school teachers having switched the municipality in which they are working as a primary school teacher at a given age, conditional on having worked as a primary school teacher in the previous year. Data for all active primary school teachers in the years 1990-2000. (back)

Table A1: First stage in teacher quality sample

Total share 60	0.067
	(0.043)
Total share 60 * post	–0.223***
	(0.050)
M*Post FE	X
Y FE	
R*Y FE	X
M*Y controls	
Adj. R^2	0.870
Obs	4351
Dep mean	.34

Note: Estimates for Equation 7 for restricted sample of municipalities where at least one teacher test score is observable at baseline. Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted. * p<0.1, ** p<0.05, *** p<0.01 (back)

Complementarities between male and female teachers

Table A2: Complementarities: Low share male - Main outcomes at age 25

	Employed/ Student	Unemployed	DI/ Pension	Other out of LF
Low: Avg share male	0.752	–0.149	–0.199	–0.344
	(0.599)	(0.342)	(0.197)	(0.294)
Municipal FE	X	X	X	X
Region*Cohort FE	X	X	X	X
Ind. controls	X	X	X	X
MOP F^{eff}	3.78	3.78	3.78	3.78
Obs	579,101	579,101	579,101	579,101
Dep mean	.846	.082	.017	.053
Std effect	.099	-.026	-.073	-.073

1.05

Table A3: Complementarities: High share male - Main outcomes at age 25

	Employed/ Student	Unemployed	DI/ Pension	Other out of LF
High: Avg share male	0.663	0.002	–0.208	–0.453
	(0.488)	(0.301)	(0.161)	(0.280)
Municipal FE	X	X	X	X
Region*Cohort FE	X	X	X	X
Ind. controls	X	X	X	X
MOP F^{eff}	5.00	5.00	5.00	5.00
Obs	226,794	226,794	226,794	226,794
Dep mean	.832	.097	.018	.052
Std effect	.098	0	-.088	-.113

1.05

Note: IV estimates for Equation 9, split sample by initial share male teachers in a municipality in 1990. Outcomes are pupils' labor market status (c.f. Table 5). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table A4: Teachers having children in municipalities

1.05

	Birth total	Birth fem	Birth male	Maternity leave
Share 60	–0.127***	–0.109***	–0.018	–0.003
	(0.040)	(0.029)	(0.025)	(0.030)
Share 60 * post	0.075	0.036	0.039	–0.040
	(0.050)	(0.036)	(0.029)	(0.032)
Y FE	X	X	X	X
Adj. R^2	0.007	0.006	0.012	0.005
Obs	4448	4448	4448	4448
Dep mean	.06	.03	.02	.02

Note: Estimates for Equation 8. Outcomes from left to right are: Share of teachers with the birth of a child, share of teachers with a birth who are female (column 2) and male (column 3), share of teachers who are female and on a leave after birth (defines as not being an active teacher in a given year after having given birth). Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted. * p<0.1, ** p<0.05, *** p<0.01 (back)

IV estimates by gender

Table A5: By gender: Labor market outcomes at age 25

	Employed/ Student	Unemployed	DI/ Pension	Other out of LF
Boys * avg share male	0.685***	–0.007	–0.145*	–0.526***
	(0.254)	(0.161)	(0.078)	(0.143)
Girls * avg share male	0.320	–0.079	–0.102	–0.099
	(0.243)	(0.156)	(0.077)	(0.139)
1.05 Municipal FE	X	X	X	X
Region*Cohort FE	X	X	X	X
Ind. controls	X	X	X	X
MOP F^{eff}	15.37	15.37	15.37	15.37
Obs	811,392	811,392	811,392	811,392
Boys: Dep mean	.84	.102	.019	.037
Girls: Dep mean	.845	.07	.015	.07
Boys: Std effect	.119	-.001	-.067	-.179
Girls: Std effect	.056	-.02	-.053	-.025

Note: Estimates for Equation 9. Outcomes are mutually exclusive categories of pupils' labor market status measured at age 25: Being in employment or a student, unemployed, on disability insurance (DI) or receiving pension payments, or being out of the labor force for other reasons. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table A6: By gender: Highest degree achieved at age 25

	Compulsory	Vocational			Academic		
	schooling	Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Boys * avg share	0.242	–0.194	0.458**	–0.262	–0.432*	0.317**	–0.128
male	(0.176)	(0.277)	(0.216)	(0.217)	(0.232)	(0.143)	(0.091)
Girls * avg	–0.632***	0.092	0.391*	0.133	–0.446*	0.465***	–0.003
1.05 share male	(0.185)	(0.274)	(0.208)	(0.219)	(0.236)	(0.156)	(0.101)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.36	15.36	15.36	15.36	15.36	15.36	15.36
Obs	810,065	810,065	810,065	810,065	810,065	810,065	810,065
Boys: Dep mean	.152	.378	.081	.094	.231	.042	.022
Girls: Dep mean	.101	.251	.136	.201	.19	.067	.054
Boys: Std effect	.043	-.025	.107	-.057	-.065	.101	-.056
Girls: Std effect	-.133	.013	.073	.021	-.072	.118	-.001

Note: IV estimates of Equation 9. Outcomes are mutually exclusive categories of pupils' highest degree achieved at age 25: Having only Compulsory education. For the Vocational track: Having a basic three year secondary degree (Sec), having additional qualifications or high school coursework beyond a basic degree (Sec Plus), having a tertiary degree from a polytechnic (Tert). For the Academic track: Having a three year high school degree (Sec), having a three year university BA degree (Tert: BA), having a two year university MA degree (Tert: MA) or higher. Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table A7: By gender: Applications and enrollment for post-compulsory education

	Apply directly	Apply late	Apply never	Pref. choice	No spot	Enrolled at 16	Enrolled ever
Boys * avg share male	0.478**	–0.405**	–0.073	0.585**	–0.207**	0.720**	0.110
	(0.205)	(0.184)	(0.076)	(0.254)	(0.104)	(0.323)	(0.079)
Girls * avg share male	0.364*	–0.294*	–0.070	0.498**	–0.073	0.474	0.141*
1.05	(0.193)	(0.176)	(0.072)	(0.247)	(0.099)	(0.301)	(0.074)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.28	15.28	15.28	15.28	15.28	13.00	13.00
Obs	825,094	825,094	825,094	825,094	825,094	695,340	695,340
Boys: Dep mean	.889	.086	.025	.857	.041	.845	.977
Girls: Dep mean	.933	.047	.02	.867	.039	.876	.982
Boys: Std effect	.097	-.092	-.03	.106	-.067	.124	.046
Girls: Std effect	.093	-.088	-.032	.093	-.024	.09	.066

Note: IV estimates for Equation 9. Outcomes in columns 1-3 are mutually exclusive categories of applications to upper secondary education: Pupils apply directly in spring of the year in which they turn 16 (Apply directly), they apply up to four years after they have turned 16 (Apply late), or they apply never or later than five years after having turned 16 (Apply never). "Allocation" (columns 4-5): Pupils obtain one of their first two preferred choices in the application (Pref. choice), or do not obtain a study slot (No spot). "Enrollment" (columns 6-7): Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at age 16), and ever enrolled in upper secondary education up to age 25 (Ever enrolled). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table A8: By gender: Aspirations for post-compulsory education

	Apply	Choose:		Get:		
	never	any Voc	only Acad	no spot	Voc	Acad
Boys * avg share	–0.073	0.811**	–0.744**	–0.207**	0.263	0.017
male	(0.076)	(0.381)	(0.375)	(0.104)	(0.324)	(0.350)
Girls * avg	–0.070	–0.367	0.431	–0.073	–0.820**	0.962***
share male	(0.072)	(0.402)	(0.395)	(0.099)	(0.348)	(0.371)
1.05 Municipal FE	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X
MOP F^{eff}	15.28	15.28	15.28	15.28	15.28	15.28
Obs	825,094	825,094	825,094	825,094	825,094	825,094
Boys: Dep mean	.025	.629	.346	.041	.501	.433
Girls: Dep mean	.02	.443	.537	.039	.329	.612
Boys: Std effect	-.03	.107	-.1	-.067	.034	.002
Girls: Std effect	-.032	-.047	.055	-.024	-.111	.126

Note: IV estimates for Equation 9. Outcomes are mutually exclusive categories for columns 1-3: Pupils 'Apply Never', pupils put in a vocational choice in any field (Choose any Voc), or pupils put in only academic track choices (Choose only Acad) (We don't report an estimate for the separate category of 287 pupils who never put in a choice, but obtain a study slot nevertheless). Columns 1 and 4-6 are also mutually exclusive categories: Pupils 'Apply never', get allocated a spot in a vocational track (Voc), or get allocated a spot in the academic track (Acad). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Table A9: Boys: Field of education at age 25

	Male	Neutral	Female	STEM	STEM-M	Education/ Teacher	Primary Teacher
Boys: Avg share male	0.506	−0.679*	0.172	0.653	0.492	0.032	0.032
1.05	(0.388)	(0.408)	(0.188)	(0.439)	(0.425)	(0.060)	(0.046)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	14.91	14.91	14.91	14.91	14.91	14.91	14.91
Obs	415,571	415,571	415,571	415,571	415,571	415,571	415,571
Dep mean	.526	.39	.084	.412	.446	.009	.005
Std effect	.065	-.089	.04	.085	.063	.022	.031

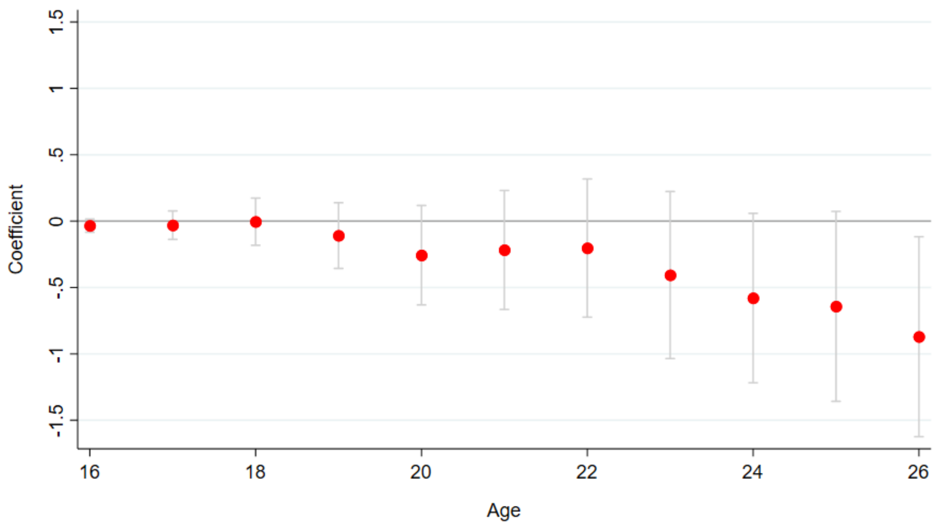
Table A10: Girls: Field of education at age 25

	Male	Neutral	Female	STEM	STEM-M	Education/ Teacher	Primary Teacher
Girls: Avg share male	0.089	−0.340	0.252	0.550**	0.977**	−0.049	0.101
1.05	(0.212)	(0.360)	(0.321)	(0.264)	(0.399)	(0.128)	(0.088)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.75	15.75	15.75	15.75	15.75	15.75	15.75
Obs	395,821	395,821	395,821	395,821	395,821	395,821	395,821
Dep mean	.069	.478	.453	.108	.309	.038	.017
Std effect	.022	-.043	.032	.113	.134	-.016	.05

Note: IV estimates of Equation 9, separate regressions by gender. Outcomes from left to right: Field is 'Male' dominated ($\geq$ 40% male), (gender) 'Neutral' or 'Female' dominated ($\geq$ 40% female), based on previous generation. Field is STEM or STEM + Medicine (STEM-M). Field is Education Science or Teacher. Field is Primary School Teacher. Individual level controls are measured at age 7 and include language (SE/FI), foreign origin, single parent HH, highest degree attained in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

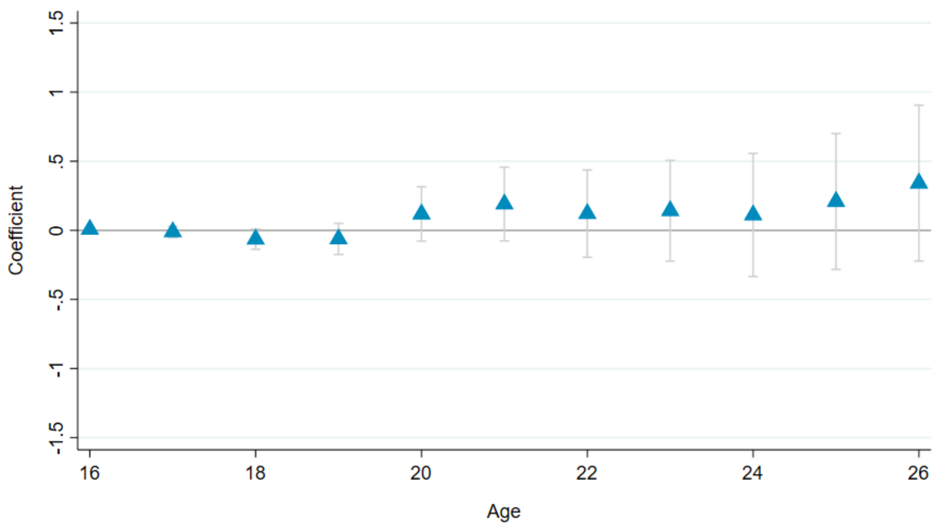
Fertility

Figure A2: Female pupils: Probability of first birth having occurred by age



Note: IV estimates of Equation 9. Outcome: First birth having occurred by age. Individual level controls are measured at age 7 and include language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. (back)

Figure A3: Male pupils: Probability of first birth having occurred by age



Note: IV estimates of Equation 9. Outcome: First birth having occurred by age. Individual level controls are measured at age 7 and include language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. (back)

Table A11: Female pupils: Probability of first birth having occurred by age

	16	17	18	19	20	21	22	23	24	25	26
Avg share male	–0.034	–0.031	–0.004	–0.109	–0.257	–0.217	–0.203	–0.406	–0.579*	–0.642*	–0.871**
	(0.026)	(0.055)	(0.091)	(0.126)	(0.191)	(0.229)	(0.266)	(0.321)	(0.326)	(0.365)	(0.384)
1.05 Municipal FE	X	X	X	X	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X	X	X	X	X
MOP F^{eff}	15.75	15.75	15.75	15.75	15.75	15.75	15.75	15.75	15.75	15.75	15.75
Obs	396,108	396,108	396,108	396,108	396,108	396,108	396,108	396,108	396,108	396,108	396,108
Dep mean	.001	.005	.013	.03	.056	.087	.119	.152	.188	.229	.273
Std effect	-.059	-.028	-.002	-.041	-.071	-.049	-.04	-.072	-.094	-.097	-.124

Table A12: Male pupils: Probability of first birth having occurred by age

Boys: Has a child by age											
	16	17	18	19	20	21	22	23	24	25	26
Avg share male	0.007	–0.011	–0.064*	–0.063	0.119	0.190	0.121	0.142	0.111	0.209	0.342
	(0.009)	(0.021)	(0.037)	(0.057)	(0.100)	(0.136)	(0.161)	(0.186)	(0.227)	(0.251)	(0.288)
Municipal FE	X	X	X	X	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X	X	X	X	X
MOP F^{eff}	14.92	14.92	14.92	14.92	14.92	14.92	14.92	14.92	14.92	14.92	14.92
Obs	415,679	415,679	415,679	415,679	415,679	415,679	415,679	415,679	415,679	415,679	415,679
Dep mean	0	.001	.003	.008	.018	.032	.052	.075	.101	.131	.165
Std effect	.027	-.023	-.076	-.045	.057	.069	.035	.034	.023	.039	.059

Note: IV estimates for Equation 9. Outcome for male pupils is the likelihood of having had the first birth at each age, from 16 to 26. Individual level controls are measured at age 7 and include language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

Results by single parent status

Table A13: By single parent status: Labor market outcomes at age 25

	Employed/ Student	Unemployed	DI/ Pension	Other out of LF
Both: Avg share male	0.461*	–0.012	–0.107	–0.321**
	(0.239)	(0.153)	(0.074)	(0.138)
Single: Avg share male	0.986***	–0.272	–0.315***	–0.354*
	(0.349)	(0.266)	(0.114)	(0.187)
Municipal FE	X	X	X	X
Region*Cohort FE	X	X	X	X
Ind. controls	X	X	X	X
MOP F^{eff}	15.32	15.32	15.32	15.32
Obs	804,799	804,799	804,799	804,799
Both: Dep mean	.854	.08	.016	.049
Single: Dep mean	.776	.125	.024	.075
Both: Std effect	.084	-.003	-.055	-.096
Single: Std effect	.138	-.048	-.121	-.079

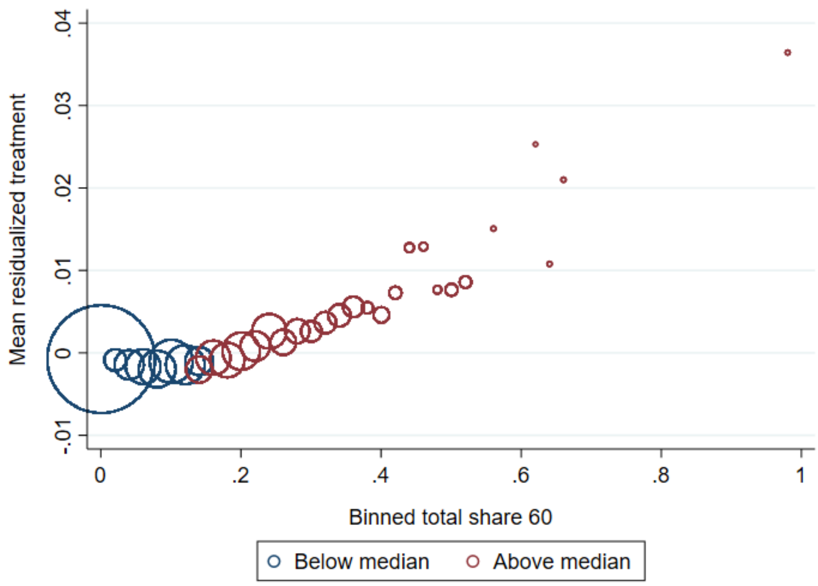
Table A14: By single parent status: Highest degree achieved at age 25

	Compulsory	Vocational			Academic		
	schooling	Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Both: Avg share male	–0.226	0.025	0.384*	–0.093	–0.352	0.319**	–0.056
	(0.155)	(0.263)	(0.205)	(0.211)	(0.228)	(0.145)	(0.093)
Single: Avg share male	0.233	–0.591	0.741***	0.077	–1.258***	1.027***	–0.229*
	(0.249)	(0.370)	(0.264)	(0.268)	(0.328)	(0.233)	(0.119)
Municipal FE	X	X	X	X	X	X	X
Region*Cohort FE	X	X	X	X	X	X	X
Ind. controls	X	X	X	X	X	X	X
MOP F^{eff}	15.31	15.31	15.31	15.31	15.31	15.31	15.31
Obs	803,494	803,494	803,494	803,494	803,494	803,494	803,494
Both: Dep mean	.11	.309	.11	.155	.217	.058	.041
Single: Dep mean	.227	.362	.097	.093	.171	.033	.018
Both: Std effect	-.047	.003	.079	-.017	-.055	.088	-.018
Single: Std effect	.032	-.072	.146	.015	-.195	.336	-.101

Note: IV estimates for Equation 9. Heterogeneity with respect to whether pupils live with two parents (Both) or a single parent (Single) at age 7. Outcomes are pupils' labor market status (c.f. Table 5) and highest degree achieved (c.f. Table 7). Individual level controls are measured at age 7 and include gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level. * p<0.1, ** p<0.05, *** p<0.01 (back)

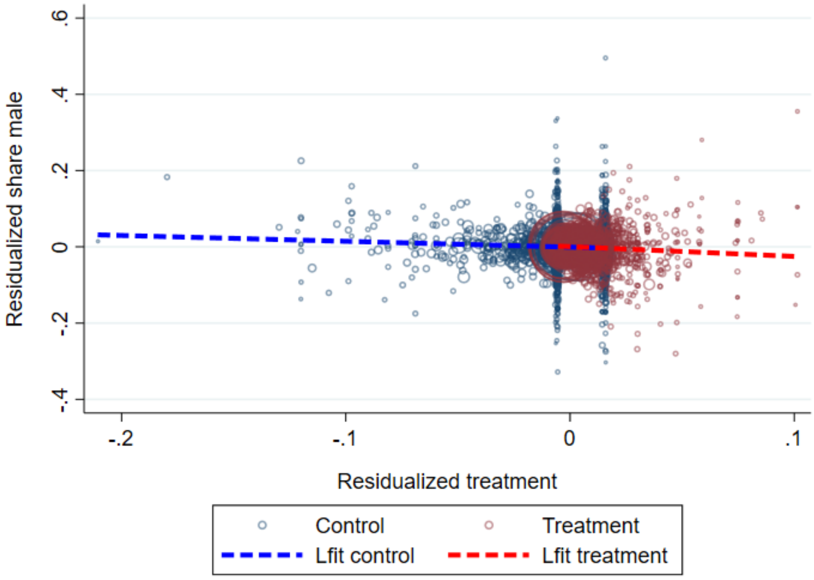
TWFE robustness

Figure A4: Municipality level: Residuals by treatment intensity



Note: Residuals of Equation 12. Average residual against binned total exposure to retirements in the post period (treatment variable). Weighted by number of municipalities per bin. (back)

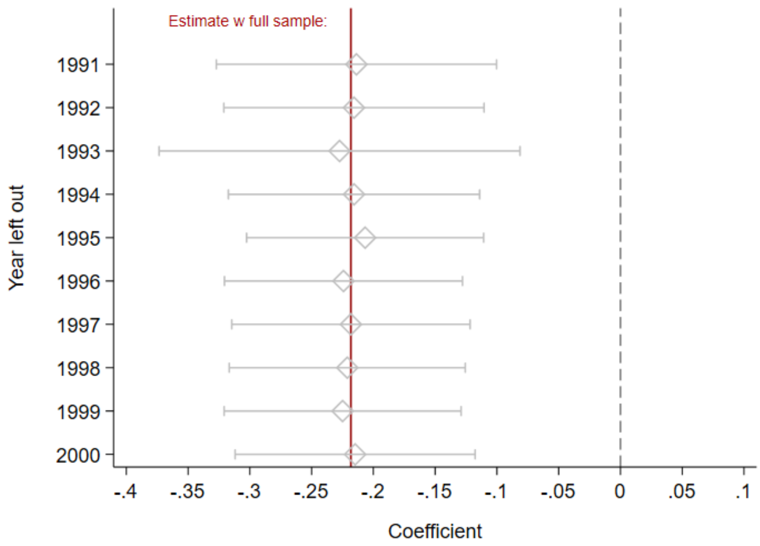
Figure A5: Municipality level: Residual plot



Note: Residual plot of Equation 7 with residualized outcome against residualized treatment. Data and linear fits are population weighted. Each dot is a municipal*year observation. Control observation are those without a retirement in the post-quota period in that municipal*year cell, and Treatment observations are those with a positive retirement value in the post-quota period in that municipal*year cell. (back)

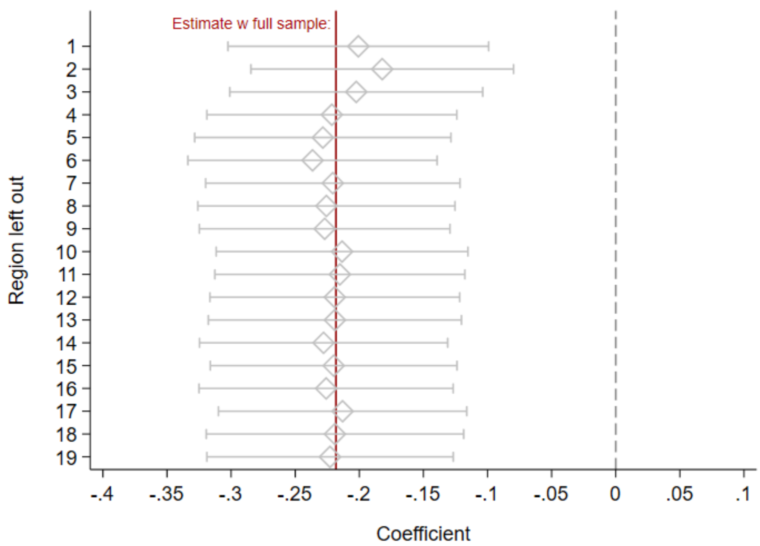
Leave one out estimation

Figure A6: Municipality level: Leave one out - years



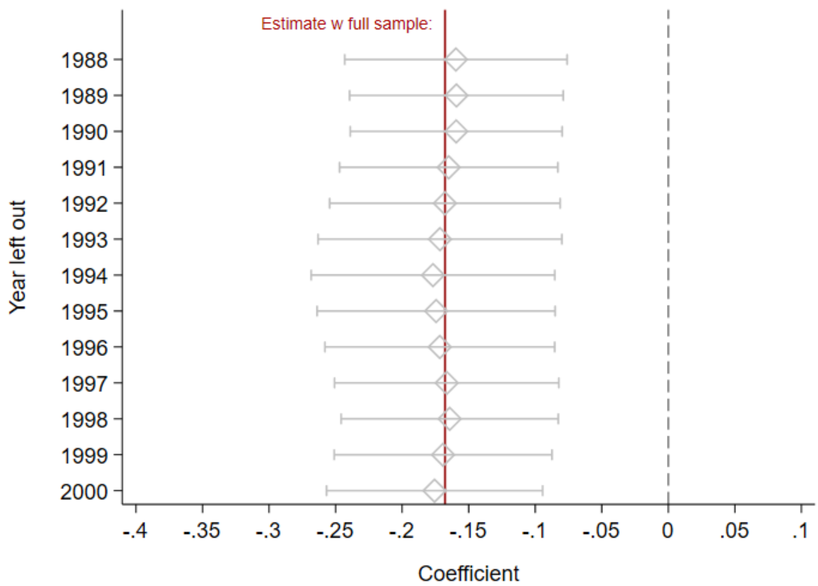
Note: Leave one out estimation of treatment coefficient in Equation 7 with respect to years. (back)

Figure A7: Municipality level: Leave one out - region



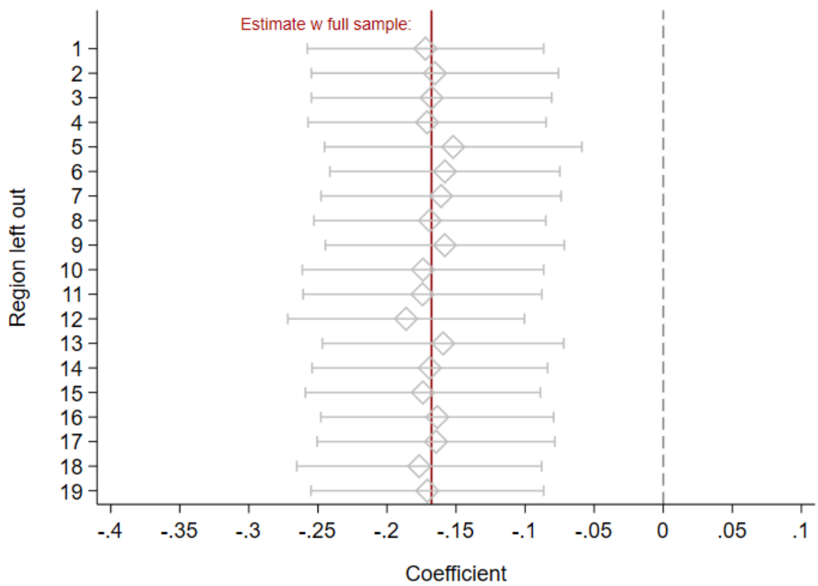
Note: Leave one out estimation of treatment coefficient in Equation 7 with respect to regions. (back)

Figure A8: Pupil level: Leave one out - cohort



Note: Leave one out estimation of treatment coefficient in Equation 11 with respect to cohorts.

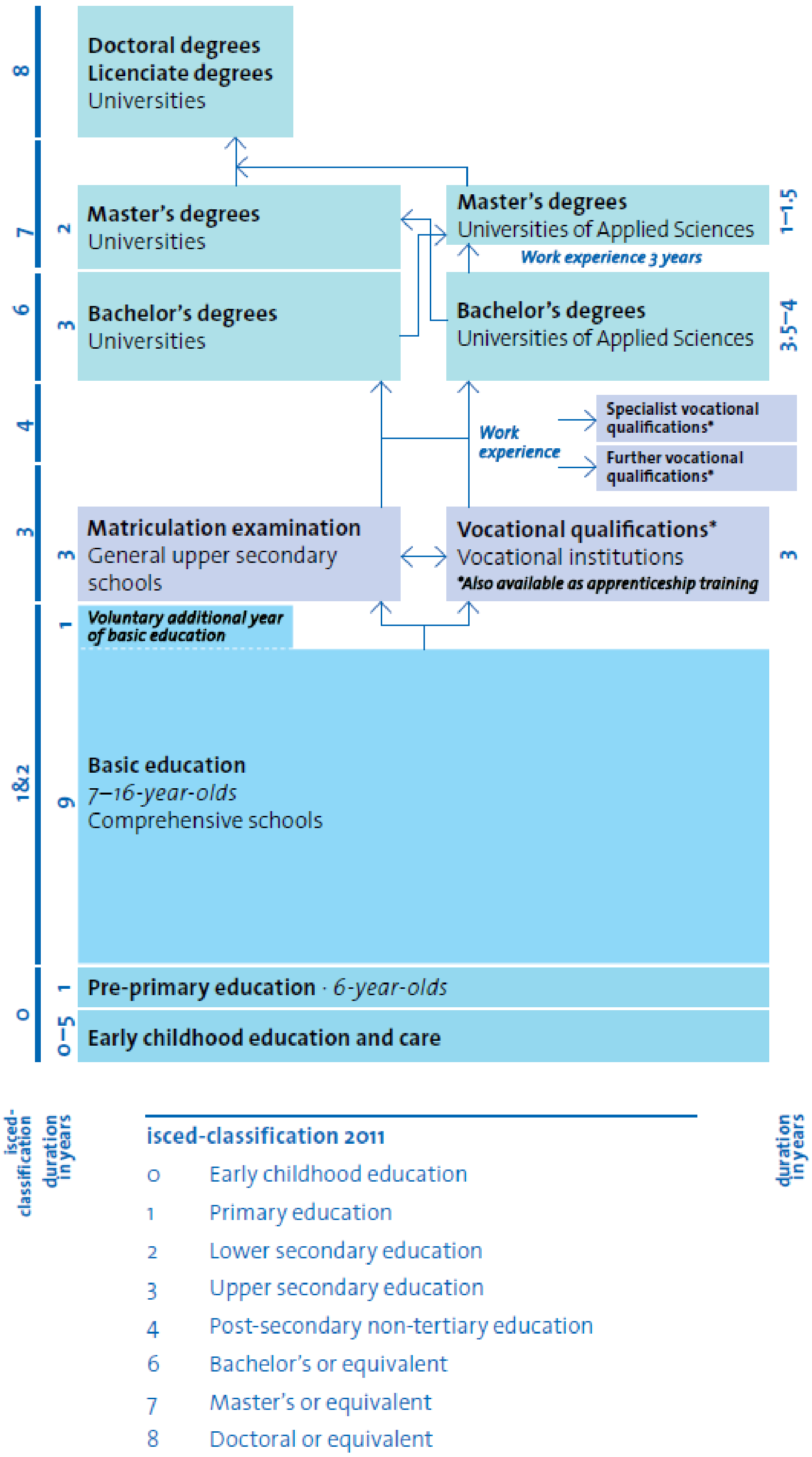
Figure A9: Pupil level: Leave one out - region



Note: Leave one out estimation of treatment coefficient in Equation 11 with respect to regions.
(back)

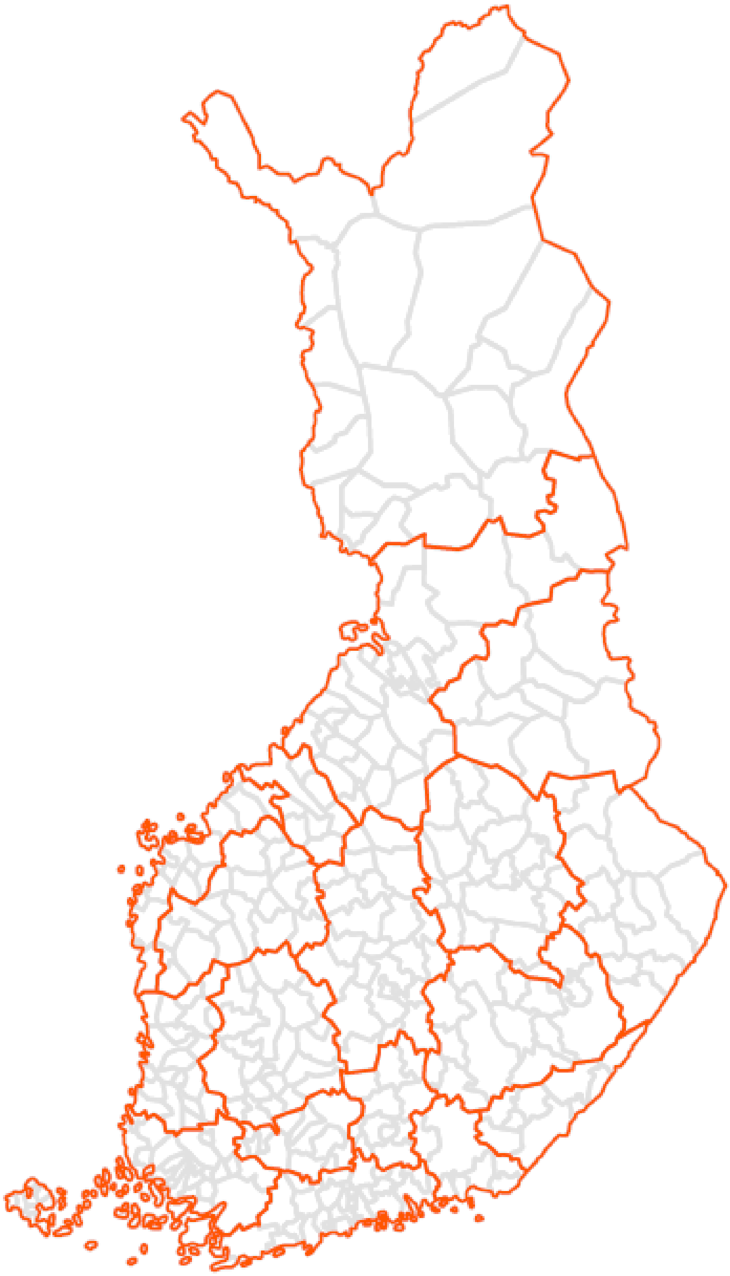
Context

Figure A10: Finnish Education System



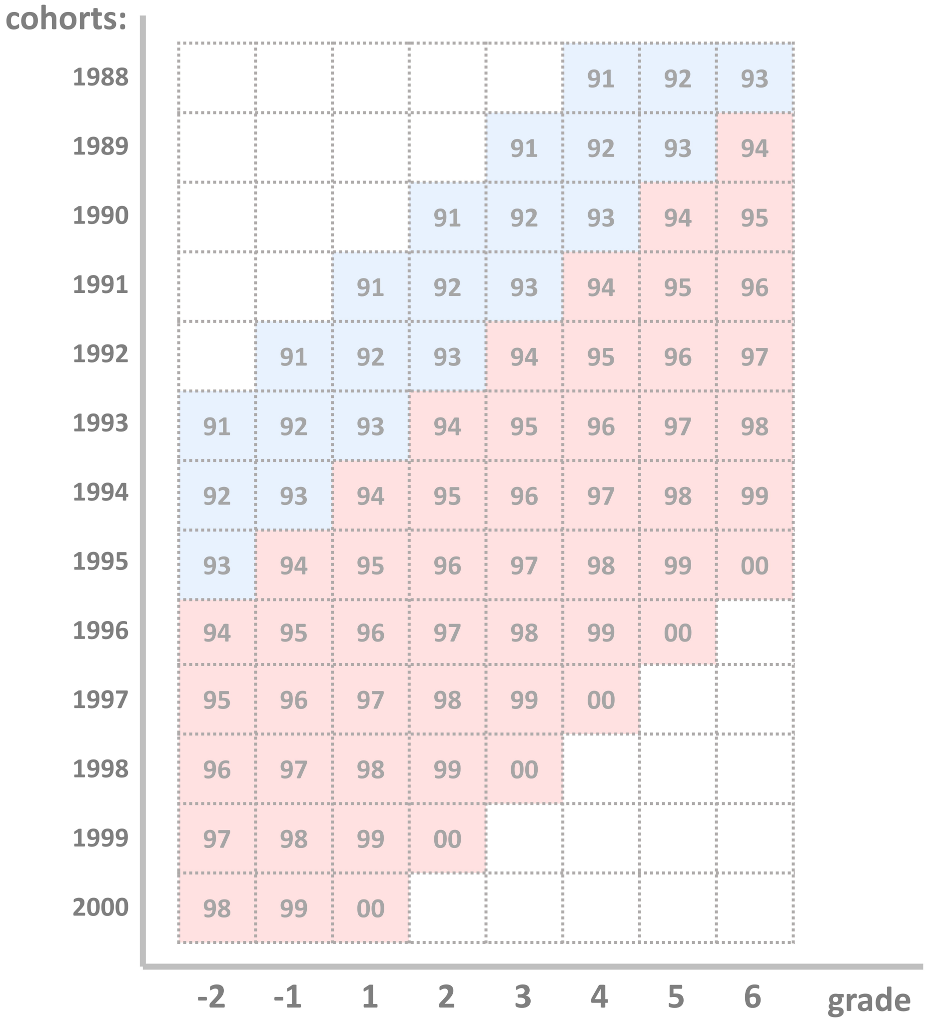
Note: Source: Ministry of Education, Finland. (back)

Figure A11: Region and municipality borders, Finland



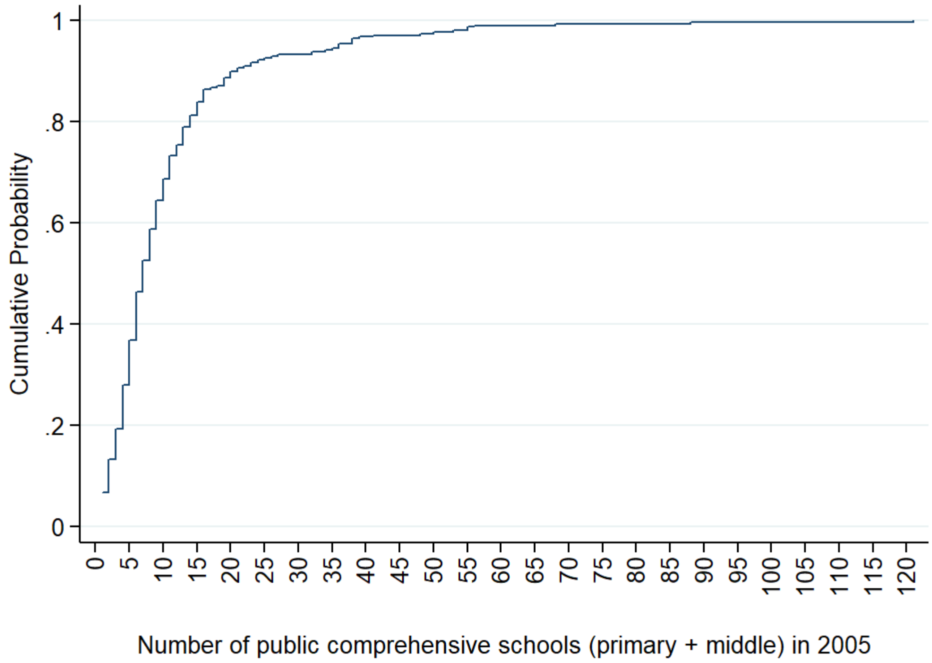
Note: Borders for 2019, shapefiles provided by Statistics Finland.

Figure A12: Pupil cohorts and exposure to quota period



Note: Figure shows cohorts by year in which they turn seven years old, and exposure to the quota by the grades which they spend in primary school. Years in which the quota was still in place colored in blue, years in which the quota was abolished in red. (back)

Figure A13: CDF number of comprehensive (primary + middle) schools by municipality, 2005



Note: Figure shows CDF of number of total comprehensive schools by municipality. Not possible to differentiate by middle schools and primary schools. Data for 2005. (back)