Supplemental Appendix

Quota vs. Quality?
Long-Term Gains from an Unusual Gender Quota

by Schaede & Mankki

Table of Contents

A Summary Statistics **2**

B Conceptual Framework: Appendix **7**

C First Stage: Teachers' Matriculation Exam Scores **10**

D Additional IV Estimates **11**

E IV Estimates by Pupil Gender: Main Outcomes **14**

F Reduced Form Estimates **19**

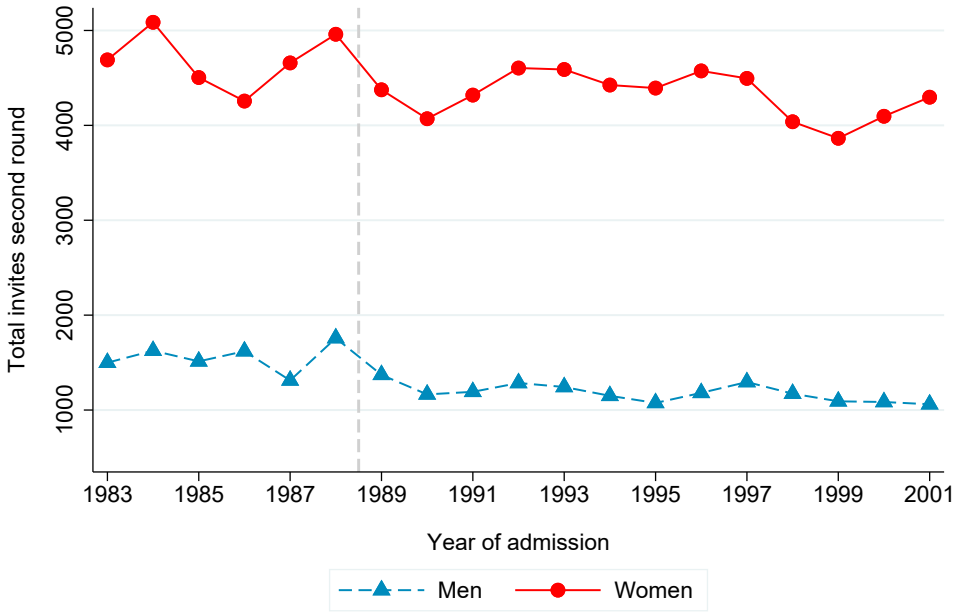
G Mechanism **23**

H Robustness **32**

I Context **42**

A Summary Statistics

Figure A1: Total Applications by Gender



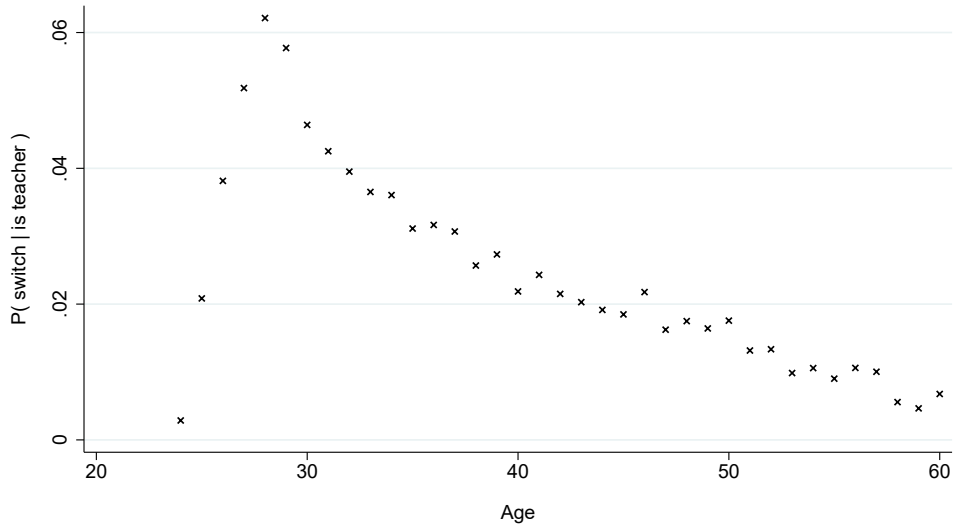
Note: Total number of male and female applicants to primary teacher studies. *Source:* Liimatainen (2002).

Table A1: Summary Statistics of Teachers by Gender

	Quota			Post-Quota		
Variable	Female	Male	Difference	Female	Male	Difference
A. Background and current place of living						
Urban residence at birth	0.586	0.563	-0.023***	0.613	0.590	-0.022***
			(0.007)			(0.007)
Rural residence at birth	0.371	0.391	0.021**	0.359	0.376	0.017**
			(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)
Total obs	14,999	7,303		18,075	7,887	
B. Education path (born after 1952)						
High school degree	0.980	0.978	-0.002	0.978	0.971	-0.007*
			(0.003)			(0.003)
Teaching degree	0.897	0.891	-0.006	0.905	0.886	-0.018***
			(0.006)			(0.005)
Primary teaching degree	0.845	0.870	0.025**	0.833	0.855	0.022***
			(0.007)			(0.006)
Age at high school degree	19.23	19.47	0.24***	19.19	19.40	0.21***
			(0.02)			(0.01)
Age at teaching degree	26.45	27.04	0.59***	26.84	27.51	0.67***
			(0.12)			(0.09)
Age at primary teacher degree	25.21	26.37	1.16***	25.87	26.88	1.01***
			(0.08)			(0.07)
C. Academic performance (born after 1952)						
Matriculation exam	0.986	0.983	-0.003	0.982	0.976	-0.006*
			(0.003)			(0.002)
Overall GPA, first take	4.78	4.22	-0.56***	4.72	4.22	-0.49***
			(0.02)			(0.01)
Overall GPA, best take	4.85	4.35	-0.51***	4.79	4.35	-0.45***
			(0.02)			(0.01)
Mathematics exam	0.741	0.826	0.086***	0.768	0.842	0.074***
			(0.009)			(0.007)
Advanced mathematics exam	0.281	0.387	0.106***	0.268	0.390	0.122***
			(0.010)			(0.008)
Total obs	7,056	3,274		11,701	4,614	

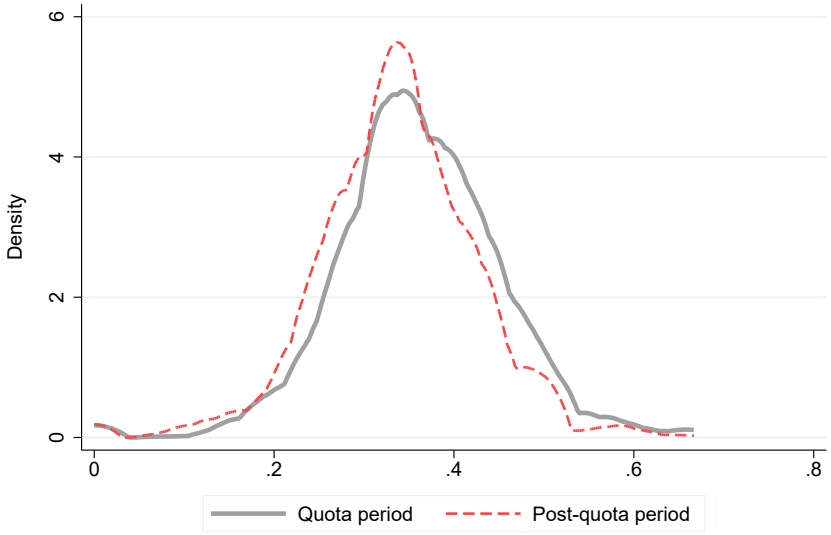
Note: Characteristics of male and female primary school teachers who are active teachers for at least one year in the quota period (1990-93) or in the post-quota period (1994-2000) and who are between 24 and 60 years old.

Figure A2: Probability of Switching Municipality of Work for Active Teachers by Age



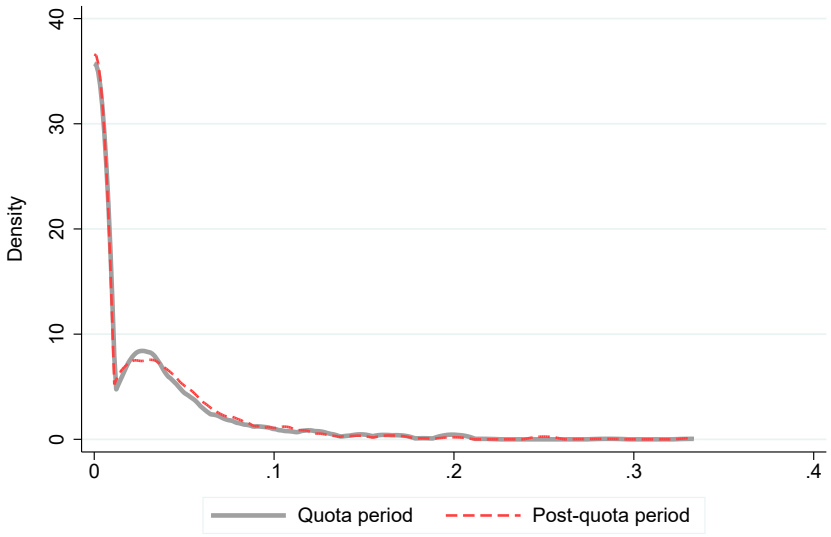
Note: Share of primary school teachers having switched the municipality in which they are working as a primary school teacher at a given age, conditional on having worked as a primary school teacher in the previous year. Data for all active primary school teachers in the years 1990-2000.

Figure A3: Distribution of Share Male Primary Teachers



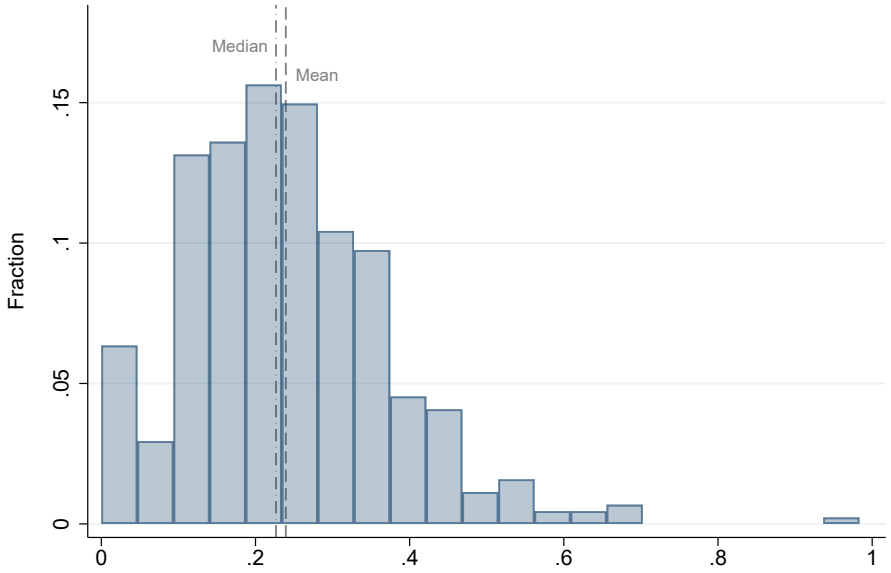
Note: Smoothed density of yearly share of male primary teachers at the municipal level, separately by years in the quota period (1991-93) and post-quota period (1994-2000).

Figure A4: Distribution of Share Primary Teachers Turning 60



Note: Smoothed density of yearly municipal share of primary school teachers turning 60, separately by years in the quota period (1991-93) and post-quota period (1994-2000).

Figure A5: Distribution of Total Share Primary Teachers Turning 60, 1991-2000



Note: Histogram of cumulative municipal share of primary school teachers turning 60 (adding up all retirements within municipality from 1991-2000).

Table A2: Summary Statistics for Main Treatment and Outcome Variables

Variable	Mean	SD
Treatment		
Municipality Level		
Share Male	0.344	0.090
Share Teachers Turning 60	0.024	0.039
Share Male, 1990	0.359	0.105
Share Male, 2000s	0.309	0.082
Pupil Level		
Share Male	0.313	0.064
Total Share Teachers Turning 60	0.078	0.056
Outcomes		
Applications and Enrollment		
Apply	0.911	0.285
Apply Late	0.066	0.249
Apply Never	0.023	0.150
Top Choice	0.858	0.349
Enrolled at 16	0.861	0.346
Enrolled Ever	0.980	0.142
Labor Market		
Employed/Student	0.842	0.364
Unemployed	0.086	0.281
DI/Pension	0.017	0.130
Other out of LF	0.053	0.223
Highest Degree		
Compulsory Schooling	0.127	0.333
Vocational Sec	0.316	0.465
Vocational Sec Plus	0.108	0.310
Vocational Tert	0.146	0.353
Academic Sec	0.211	0.408
Academic Tert: BA	0.054	0.227
Academic Tert: MA	0.038	0.190
Field of Study		
Male	0.303	0.460
Neutral	0.433	0.495
Female	0.264	0.441
STEM	0.264	0.441
STEM-M	0.379	0.485
Education/Teacher	0.023	0.150
Primary Teacher	0.011	0.102

Note: Summary statistics for main treatment and outcome variables. Municipality Level refers to the municipal-by-year panel, while Pupil Level refers to exposure measures throughout primary school for the pupil panel.

B Conceptual Framework: Appendix

B.1 Varying Signal Precision

In the Finnish context, as documented in Mechanism Section VI, the matriculation exam score is not predictive of teacher performance, i.e., a quite noisy signal of ability. In contrast, in US college admissions, standardized test scores have been shown to be predictive of students' academic success in the first year of college (Friedman et al., 2025). We therefore might want to consider how the impacts of requiring a group-blind admissions regime differ as the test score signal becomes more precise when a test is biased. To examine comparative statics with respect to signal precision, let

$$\Upsilon_j = \{1 - \Phi(z^*_M)\}\ \sigma_a\, \rho\, \lambda(z^*_M) + \{1 - \Phi(z^*_F)\}\ \sigma_a\, \rho\, \lambda(z^*_F) \tag{13}$$

denote the total expected ability of admits. For ease of exposition, we set $\mu_a = 0$.

B.1.1 Unconstrained Admissions

When an admissions office is unconstrained, optimal threshold scores take into account any group-specific bias (Equation 4): $s^*_M = s^*_F - b_M$, such that:

$$z^*_M = z^*_F = \frac{\rho}{\sigma_a}\, s^*_F$$

The capacity constraint and total expected ability of admits are, respectively:

$$2\,(1 - \Phi(z^*_F)) = c \qquad\qquad \Upsilon = 2\,(1 - \Phi(z^*_F))\,\rho\,\sigma_a\,\lambda(z^*_F) = 2\,\rho\,\sigma_a\,\phi(z^*_F)$$

Representation: Taking the derivative of the capacity constraint with respect to ρ, we obtain:

$$-2\,\phi(z^*_F)\,\frac{dz^*_F}{d\rho} = 0$$

Since $\phi(\cdot) > 0$, it must be that $\dfrac{dz^*_F}{d\rho} = 0$, and thus the mass of admitted male and female candidates does not change when increasing the precision of the test.

Total Expected Ability of Admits: Taking the derivative of Υ with respect to ρ, we obtain:

$$2\,\sigma_a\,\phi(z^*_F) + 2\,\rho\,\sigma_a\,\phi'(z^*_F)\,\frac{dz^*_F}{d\rho} > 0$$

since $\phi(\cdot) > 0$ and $\dfrac{dz^*_F}{d\rho}$ is equal zero. Total expected ability for both groups therefore increases when precision of the test increases. In sum, when an admissions office is unconstrained, a test with a more precise signal leaves representation unchanged and selects candidates with higher levels of expected ability.

B.1.2 Group-blind Admissions

In contrast, when admissions are required to be group-blind, a more precise signal under-admits the group against whom the test is biased and does not necessarily deliver efficiency gains. Under group-blind admissions, we have that $s^* = s_M^* = s_F^*$ and $z_M^* > z_F^*$.

Representation: Differentiating the capacity constraint $(1 - \Phi(z_M^*)) + (1 - \Phi(z_F^*)) = c$ with respect to ρ, we obtain:

$$-\phi(z_M^*)\left(\frac{z_M^*}{\rho} + \frac{ds^*}{d\rho} * \frac{\rho}{\sigma_a}\right) = \phi(z_F^*)\left(\frac{z_F^*}{\rho} + \frac{ds^*}{d\rho} * \frac{\rho}{\sigma_a}\right) \tag{14}$$

We can assess in which direction the score cut-off moves when precision increases by rearranging the above Equation 14:

$$\frac{ds^*}{d\rho} = -\frac{\sigma_a\left(\phi(z_M^*)\, z_M^* + \phi(z_F^*)\, z_F^*\right)}{\rho^2\left(\phi(z_M^*) + \phi(z_F^*)\right)} < 0$$

since admissions are from the right tail of the distribution. The score cut-off thus decreases as precision increases. Intuitively, a test with higher precision will decrease the variance of the score distribution, such that the tails are thinner. Admitting the same mass of candidates therefore implies a lower s^*.

We can show by contradiction that the mass of men (LHS of Equation 14) falls as precision increases. Note first that it is not possible for the change in the mass of admits of both groups to be zero ($\Delta N_M = -\Delta N_F = 0$), since $\phi(\cdot) > 0$ and $z_M^* > z_F^*$.
Suppose the LHS of Equation 14 were > 0, then it must be that

$$z_M^* < -\frac{ds^*}{d\rho} * \frac{\rho^2}{\sigma_a} \qquad \text{and} \qquad z_F^* > -\frac{ds^*}{d\rho} * \frac{\rho^2}{\sigma_a}$$

But this is not possible since, $z_M^* > z_F^*$. It therefore follows that the LHS of Equation 14 is < 0 (and thus $\dfrac{dz_M^*}{d\rho} > 0$), i.e. the mass of men **falls** as the precision of the test increases.

Total Expected Ability of Admits: Taking the derivative of Υ with respect to ρ and using the properties that $\dfrac{dz_F^*}{d\rho} = -\dfrac{\phi(z_M^*)}{\phi(z_F^*)}\dfrac{dz_M^*}{d\rho}$ (see Equation 14) and $\dfrac{\phi'(x)}{\phi(x)} = -x$, we obtain:

$$\begin{aligned}
&\sigma_a\,\phi(z_M^*) + \sigma_a\,\phi(z_F^*) + \rho\,\sigma_a\,\phi'(z_M^*)\,\frac{dz_M^*}{d\rho} + \rho\,\sigma_a\,\phi'(z_F^*)\,\frac{dz_F^*}{d\rho} \\
=&\,\sigma_a\,\phi(z_M^*) + \sigma_a\,\phi(z_F^*) + \rho\,\sigma_a\,\phi(z_M^*)\,\frac{dz_M^*}{d\rho}\,(z_F^* - z_M^*) \\
=&\,\sigma_a\left(\phi(z_M^*) + \phi(z_F^*)\right) - b_M\,\rho\,\sigma_a\,\phi(z_M^*)\,\frac{dz_M^*}{d\rho}
\end{aligned}$$

The first term in the above is positive, while the second term is negative, such that the net impact of increasing precision is ambiguous. Intuitively, increasing precision means getting a better signal of ability among those admitted from both groups. However, because the test exhibits a bias against men, increasing precision also incurs a cost by under-admitting candidates from the male group and over-admitting female candidates. The net effect on total expected ability of admits when selecting from a test with a more precise signal thus depends on how the gains from increased precision compare to the loss from under-admitting candidates from the disadvantaged group.

B.2 Derivations for Section II

Probability of getting admitted (Equation 5): A candidate of group g is admitted if:

$$\begin{aligned} s_g &> s_g^* \\ \frac{s_g - \mu_{s_g}}{\sigma_s} &> \frac{s_g^* - \mu_{s_g}}{\sigma_s} \\ &= \frac{\rho}{\sigma_a}\left(s_g^* - \mu_{a_g} + b_g\right) \\ &\equiv z_g^* \end{aligned}$$

The mass of candidates admitted for group g is thus equal to $1 - \Phi(z_g^*)$.

Mean expected ability among admits (Equation 6)
We can rewrite $a_g = \mu_{a_g} + \xi_{a_g}$ with $\xi_{a_g} \sim \mathcal{N}(0,\, \sigma^2_{\xi_{a_g}})$ such that $E[a_g|s_g > s_g^*] = \mu_{a_g} + E[\xi_{a_g}|s_g > s_g^*]$ and $E[\xi_{a_g}|s_g]$ a linear regression of ξ_{a_g} on s_g. Applying iterated expectations $E(\xi_{a_g}) = E_s[E(\xi_{a_g}|s)]$, we thus get that:

$$\begin{aligned} E[a_g|s_g > s_g^*] &= \mu_{a_g} + \sigma_a\,\rho\,E\left[\frac{s_g - \mu_{s_g}}{\sigma_s} \mid \frac{s_g - \mu_{s_g}}{\sigma_s} > \frac{s_g^* - \mu_{s_g}}{\sigma_s}\right] \\ &= \mu_{a_g} + \sigma_a\,\rho\,\lambda(z_g^*) \end{aligned}$$

with $z_g^* \equiv \dfrac{s_g^* - \mu_{s_g}}{\sigma_s}$ and $\lambda(\cdot) = \dfrac{\phi(\cdot)}{1 - \Phi(\cdot)}$ the inverse Mills Ratio.

C First Stage: Teachers' Matriculation Exam Scores

Table A3: First Stage: Teachers' Matriculation Exam Scores

	Overall Score	Language	Best of Math or Sciences	Math	Sciences	First Stage
Total Share 60 *	0.09	0.06	0.17	0.14	0.11	–0.22
Post-Quota	(0.25)	(0.27)	(0.28)	(0.51)	(0.33)	(0.05)
Total Share 60	–0.09	–0.07	–0.14	–0.34	–0.02	0.06
	(0.22)	(0.24)	(0.26)	(0.45)	(0.30)	(0.04)
Adj. R^2	0.84	0.86	0.82	0.81	0.82	0.87
Obs	4358	4358	4358	4347	4348	4358
Dep mean	4.52	4.45	4.73	3.72	4.49	.36

Note: Estimates for Equation 8 with average of local teachers' grades in first attempt of matriculation exam as the outcome. Overall score is the average grade obtained across the four mandatory exam fields. Language score includes grades for the three mandatory language fields in column 2 (mother tongue (FI/SE), second national language (SE/FI) and foreign language). The fourth mandatory field consists in a candidate's choice of either Mathematics or Sciences. The best score of Math or Sciences (column 3) is considered in the calculation of overall score if candidates took both fields. Math (column 4) and Sciences (column 5) report grades irrespective of whether they are counted towards overall score. Grades are assigned on a curve and range from 2-6, with 6 being the highest grade and 0 assigned for a failing exam (Kupiainen et al., 2018). For more detail regarding the matriculation exam, see Section VI.A. First stage on teacher test score sample in Column 6. Dependent variable mean is reported for the quota period. Data is available for teacher cohorts born after 1952. Sample is restricted to municipalities with at least one teacher with observed score in any year. All specifications include year and municipality-by-post-quota fixed effects. Standard errors clustered at the municipality level. Regressions weighted by population.

D Additional IV Estimates

D.1 Sensitivity Main Results

Table A4: IV Estimates: Applications and Enrollment for Post-Compulsory Education Municipal and Cohort Fixed Effects only

	Apply directly	Apply late	Apply never	Top choice	Enrolled at 16	Enrolled ever
Avg Share Male	0.698	–0.540	–0.159	0.715	1.072	0.213
	(0.236)	(0.189)	(0.083)	(0.281)	(0.385)	(0.087)
MOP F^{eff}	17.54	17.54	17.54	17.54	14.32	14.32
Obs	825,095	825,095	825,095	825,095	695,341	695,341
Dep mean	.911	.066	.023	.858	.861	.98
Std effect	.156	-.138	-.067	.13	.193	.094

Note: IV estimate of Equation 9. **Applications:** 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:** (column 4) Pupils obtain one of their first two choices in the application (Top choice). **Enrollment:** (columns 5-6) Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at 16), and ever enrolled in upper secondary education up to age 25 (Enrolled ever). All specifications include year and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A5: First Stage, Reduced Form and IV: Employed/Student at Age 25

	First Stage			RF			IV		
	Avg Share Male			Employed/Student			Employed/Student		
Avg Share Male							0.297	0.404	0.510
							(0.215)	(0.216)	(0.242)
Total Share 60 *	–0.177	–0.177	–0.169	–0.053	–0.053	–0.076			
Post	(0.042)	(0.042)	(0.043)	(0.033)	(0.033)	(0.033)			
Total Share 60	0.034	0.034	0.044	0.031	0.031	0.040	0.021	0.019	0.019
	(0.036)	(0.036)	(0.037)	(0.028)	(0.028)	(0.030)	(0.026)	(0.027)	(0.027)
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}							17.86	17.84	15.44
Adj. R^2	0.916	0.916	0.921	0.008	0.008	0.008			
Obs	811,393	811,393	811,393	811,393	811,393	811,393	811,393	811,393	811,393
Dep mean	.313	.313	.313	.842	.842	.842	.842	.842	.842

Note: Columns 1-3 show estimates for Equation 10 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 10), and Columns 7-9 show IV estimates of Equation 9 with being either employed or a student at age 25 as the outcome. Individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

D.2 Additional Results on Applications

Table A6: Obtained Choices in Post-Compulsory Applications

	First	Second	Third	Fourth	Fifth	Switch	No Spot	Apply Never
Avg Share Male	0.319	0.209	–0.136	–0.105	0.011	–0.112	–0.110	–0.075
	(0.288)	(0.161)	(0.097)	(0.066)	(0.039)	(0.090)	(0.103)	(0.073)
MOP F^{eff}	15.35	15.35	15.35	15.35	15.35	15.35	15.35	15.35
Obs	825,095	825,095	825,095	825,095	825,095	825,095	825,095	825,095
Dep mean	.773	.085	.035	.015	.007	.019	.043	.023
Std effect	.048	.048	-.047	-.055	.008	-.052	-.035	-.032

Note: IV estimates of Equation 9. Outcomes are mutually exclusive categories for allocation of slots in post-compulsory education application, from left to right: Pupils obtain their First, ..., Fifth choice. Pupils switch from assigned slot to other option (Switch), do not obtain any slot at all (No Spot), and do not put in an application within five years after middle school (Never Apply). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A7: By Gender: Aspirations for Post-Compulsory Education

	Apply	Choose:		Obtain:		
	never	any Voc	only Acad	no spot	Voc	Acad
Boys * Avg Share	–0.075	0.770	–0.701	–0.202	0.242	0.035
Male	(0.076)	(0.376)	(0.371)	(0.110)	(0.322)	(0.348)
Girls * Avg	–0.076	–0.417	0.485	–0.005	–0.895	0.976
Share Male	(0.073)	(0.399)	(0.392)	(0.104)	(0.348)	(0.368)
MOP F^{eff}	7.69	7.69	7.69	7.69	7.69	7.69
Obs	825,095	825,095	825,095	825,095	825,095	825,095
Boys: Dep mean	.025	.628	.346	.044	.498	.432
Girls: Dep mean	.02	.442	.537	.041	.327	.611
Boys: Std effect	-.031	.102	-.094	-.063	.031	.004
Girls: Std effect	-.034	-.053	.062	-.002	-.121	.127
P-value	.976	0	0	.002	0	0

Note: IV estimate of Equation 9. **Choices:** Outcomes are mutually exclusive categories for columns 1-3: Pupils 'Apply never', pupils put in a vocational degree in any of five available choices (Choose any Voc), or pupils put in only academic track choices (Choose only Acad). (We do not report an estimate for the separate category of 287 pupils who never put in a choice, but obtain a study slot nevertheless.) **Slots obtained:** Columns 1 and 4-6 are also mutually exclusive categories: Pupils 'Apply never', pupils get allocated 'no spot', a spot in a vocational track (Voc), or a spot in the academic track (Acad). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

E IV Estimates by Pupil Gender: Main Outcomes

E.1 IV Estimates by Pupil Gender

Table A8: By Gender: Applications and Enrollment for Post-Compulsory Education

	Apply directly	Apply late	Apply never	Top choice	Enrolled at 16	Enrolled ever
Boys * Avg Share	0.472	–0.397	–0.075	0.596	0.717	0.110
Male	(0.203)	(0.181)	(0.076)	(0.255)	(0.322)	(0.079)
Girls * Avg	0.359	–0.282	–0.076	0.445	0.472	0.141
Share Male	(0.191)	(0.173)	(0.073)	(0.247)	(0.300)	(0.074)
MOP F^{eff}	7.69	7.69	7.69	7.69	15.74	15.74
Obs	825,095	825,095	825,095	825,095	695,341	695,341
Boys: Dep mean	0.889	0.085	0.025	0.852	0.845	0.977
Girls: Dep mean	0.933	0.047	0.020	0.863	0.876	0.982
Boys: Std effect	.096	-.091	-.031	.107	.124	.046
Girls: Std effect	.091	-.085	-.034	.082	.09	.066
P-value	.097	.059	.976	.217	.057	.469

Note: IV estimate of Equation 9. **Applications:** 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:** (column 4) Pupils obtain one of their first two choices in the application (Top choice). **Enrollment:** (columns 5-6) Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at 16), and ever enrolled in upper secondary education up to age 25 (Enrolled ever). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A9: By Gender: Highest Degree Achieved at Age 25

	Compulsory schooling	Vocational			Academic		
		Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Boys * Avg Share	0.241	–0.192	0.456	–0.263	–0.430	0.315	–0.128
Male	(0.175)	(0.276)	(0.215)	(0.216)	(0.230)	(0.142)	(0.090)
Girls * Avg	–0.631	0.093	0.389	0.132	–0.443	0.463	–0.003
Share Male	(0.184)	(0.273)	(0.207)	(0.218)	(0.234)	(0.155)	(0.101)
MOP F^{eff}	7.73	7.73	7.73	7.73	7.73	7.73	7.73
Obs	810,066	810,066	810,066	810,066	810,066	810,066	810,066
Boys: Dep mean	0.152	0.378	0.081	0.094	0.231	0.042	0.022
Girls: Dep mean	0.101	0.251	0.136	0.201	0.190	0.067	0.054
Boys: Std effect	.043	-.025	.107	-.057	-.065	.1	-.056
Girls: Std effect	-.133	.014	.072	.021	-.072	.117	-.001
P-value	0	.111	.429	0	.891	.019	.011

Note: IV estimate 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. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A10: By Gender: Labor Market Outcomes at Age 25

	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Boys * Avg Share	0.683	–0.007	–0.144	–0.525
Male	(0.253)	(0.160)	(0.077)	(0.142)
Girls * Avg	0.319	–0.078	–0.101	–0.099
Share Male	(0.242)	(0.155)	(0.076)	(0.138)
MOP F^{eff}	7.73	7.73	7.73	7.73
Obs	811,393	811,393	811,393	811,393
Boys: Dep mean	0.840	0.102	0.019	0.037
Girls: Dep mean	0.845	0.070	0.015	0.070
Boys: Std effect	.119	-.001	-.067	-.178
Girls: Std effect	.056	-.02	-.053	-.025
P-value	.002	.395	.161	0

Note: IV estimate of 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. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A11: By Gender: Field of Education at Age 25

	Male	Neutral	Female	STEM	STEM-M	Education/ Teacher	Primary Teacher
Boys * Avg Share	0.504	–0.676	0.172	0.651	0.490	0.032	0.032
Male	(0.386)	(0.407)	(0.187)	(0.437)	(0.423)	(0.060)	(0.045)
Girls * Avg	0.088	–0.338	0.250	0.546	0.972	–0.048	0.101
Share Male	(0.211)	(0.358)	(0.319)	(0.263)	(0.397)	(0.128)	(0.088)
Boys: MOP F^{eff}	14.98	14.98	14.98	14.98	14.98	14.98	14.98
Girls: MOP F^{eff}	15.83	15.83	15.83	15.83	15.83	15.83	15.83
Boys: Obs	415,571	415,571	415,571	415,571	415,571	415,571	415,571
Girls: Obs	395,822	395,822	395,822	395,822	395,822	395,822	395,822
Boys: Dep mean	.526	.39	.084	.412	.446	.009	.005
Girls: Dep mean	.069	.478	.453	.108	.309	.038	.017
Boys: Std effect	.064	-.088	.039	.084	.063	.022	.031
Girls: Std effect	.022	-.043	.032	.112	.134	-.016	.05
P-value	.336	.507	.827	.828	.339	.552	.489

Note: IV estimates of Equation 9. Outcomes from left to right: Field is 'Male' dominated (> 70% male), (gender) 'Neutral' or 'Female' dominated (> 70% female), based on previous generation. Field is STEM or STEM + Medicine (STEM-M). Field is Education Science or Teacher. Field is Primary School Teacher. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

E.2 IV Estimates by Pupil Gender: Split Sample Estimates

Table A12: Applications and Labor Market Attachment

	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Boys * Avg Share	0.344	–0.351	0.007	0.275	–0.052	–0.078	–0.120
Male	(0.259)	(0.243)	(0.103)	(0.315)	(0.237)	(0.099)	(0.132)
Girls * Avg	0.495	–0.332	–0.163	0.724	–0.004	–0.171	–0.528
Share Male	(0.216)	(0.177)	(0.089)	(0.288)	(0.171)	(0.097)	(0.218)
Boys: MOP F^{eff}	14.90	14.90	14.90	14.98	14.98	14.98	14.98
Girls: MOP F^{eff}	15.74	15.74	15.74	15.83	15.83	15.83	15.83
Boys: Obs	421,151	421,151	421,151	415,571	415,571	415,571	415,571
Girls: Obs	403,944	403,944	403,944	395,822	395,822	395,822	395,822
Boys: Dep mean	0.889	0.085	0.025	0.840	0.102	0.019	0.037
Girls: Dep mean	0.933	0.047	0.020	0.845	0.070	0.015	0.070
P-value	0.580	0.938	0.179	0.232	0.864	0.460	0.084

Table A13: Highest Degree Achieved at Age 25

	Compulsory schooling	Vocational			Academic		
		Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Boys * Avg Share	0.179	0.062	0.208	–0.102	–0.812	0.538	–0.073
Male	(0.233)	(0.385)	(0.222)	(0.230)	(0.343)	(0.193)	(0.092)
Girls * Avg	–0.506	–0.197	0.637	–0.071	–0.028	0.233	–0.069
Share Male	(0.232)	(0.329)	(0.300)	(0.303)	(0.238)	(0.172)	(0.154)
Boys: MOP F^{eff}	14.97	14.97	14.97	14.97	14.97	14.97	14.97
Girls: MOP F^{eff}	15.82	15.82	15.82	15.82	15.82	15.82	15.82
Boys: Obs	414,896	414,896	414,896	414,896	414,896	414,896	414,896
Girls: Obs	395,170	395,170	395,170	395,170	395,170	395,170	395,170
Boys: Dep mean	0.152	0.378	0.081	0.094	0.231	0.042	0.022
Girls: Dep mean	0.101	0.251	0.136	0.201	0.190	0.067	0.054
P-value	0.053	0.601	0.195	0.926	0.038	0.174	0.978

Note: IV estimate of Equation 9, split sample estimates by gender (i.e. fully interacted model). Outcomes are pupils' applications in the last year of middle school (see Table 3) and labor market status (see Table 5). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

F Reduced Form Estimates

F.1 Reduced Form: Main Results

Table A14: Reduced Form: Applications and Enrollment for Post-Compulsory Education

	Apply directly	Apply late	Apply never	Top choice	Enrolled at 16	Enrolled ever
Total Share 60 *	–0.071	0.058	0.013	–0.089	–0.089	–0.018
Post	(0.027)	(0.024)	(0.012)	(0.036)	(0.035)	(0.010)
Total Share 60	0.017	–0.021	0.003	–0.016	0.026	0.005
	(0.021)	(0.019)	(0.009)	(0.027)	(0.028)	(0.008)
Adj. R^2	0.070	0.030	0.138	0.063	0.049	0.024
Obs	825,095	825,095	825,095	825,095	695,341	695,341
Dep mean	.911	.066	.023	.858	.861	.98

Note: Reduced Form estimates as in Equation 10. **Applications:** 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:** (column 4) Pupils obtain one of their first two choices in the application (Top choice). **Enrollment:** (columns 5-6) Pupils are enrolled in upper secondary education in the fall of the year in which they turn 16 (Enrolled at 16), and ever enrolled in upper secondary education up to age 25 (Enrolled ever). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A15: Reduced Form: Highest Degree Achieved at Age 25

	Compulsory schooling	Vocational			Academic		
		Sec	Sec Plus	Tert	Sec	Tert: BA	Tert: MA
Total Share 60 *	0.029	0.009	–0.072	0.013	0.074	–0.065	0.012
Post	(0.024)	(0.044)	(0.029)	(0.036)	(0.034)	(0.022)	(0.016)
Total Share 60	–0.007	–0.005	0.029	–0.005	–0.034	0.027	–0.005
	(0.019)	(0.036)	(0.023)	(0.028)	(0.027)	(0.016)	(0.012)
Adj. R^2	0.056	0.086	0.012	0.035	0.065	0.042	0.028
Obs	810,066	810,066	810,066	810,066	810,066	810,066	810,066
Dep mean	.127	.316	.108	.146	.211	.054	.038

Note: Reduced Form estimates as in Equation 10. 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. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

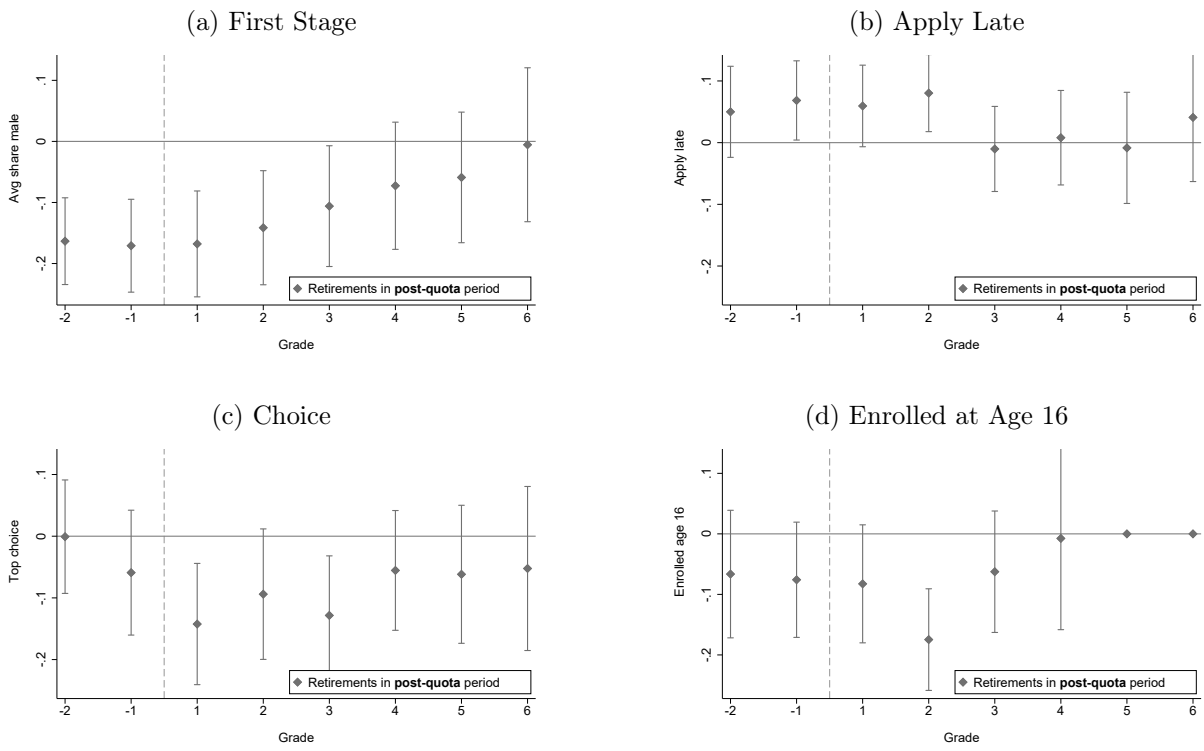
Table A16: Reduced Form: Labor Market Outcomes at Age 25

	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Total Share 60 *	–0.086	0.006	0.021	0.055
Post	(0.033)	(0.026)	(0.011)	(0.018)
Total Share 60	0.042	–0.001	–0.005	–0.035
	(0.029)	(0.022)	(0.009)	(0.015)
Adj. R^2	0.025	0.023	0.003	0.013
Obs	811,393	811,393	811,393	811,393
Dep mean	.842	.086	.017	.053

Note: Reduced Form estimates as in Equation 10. 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. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

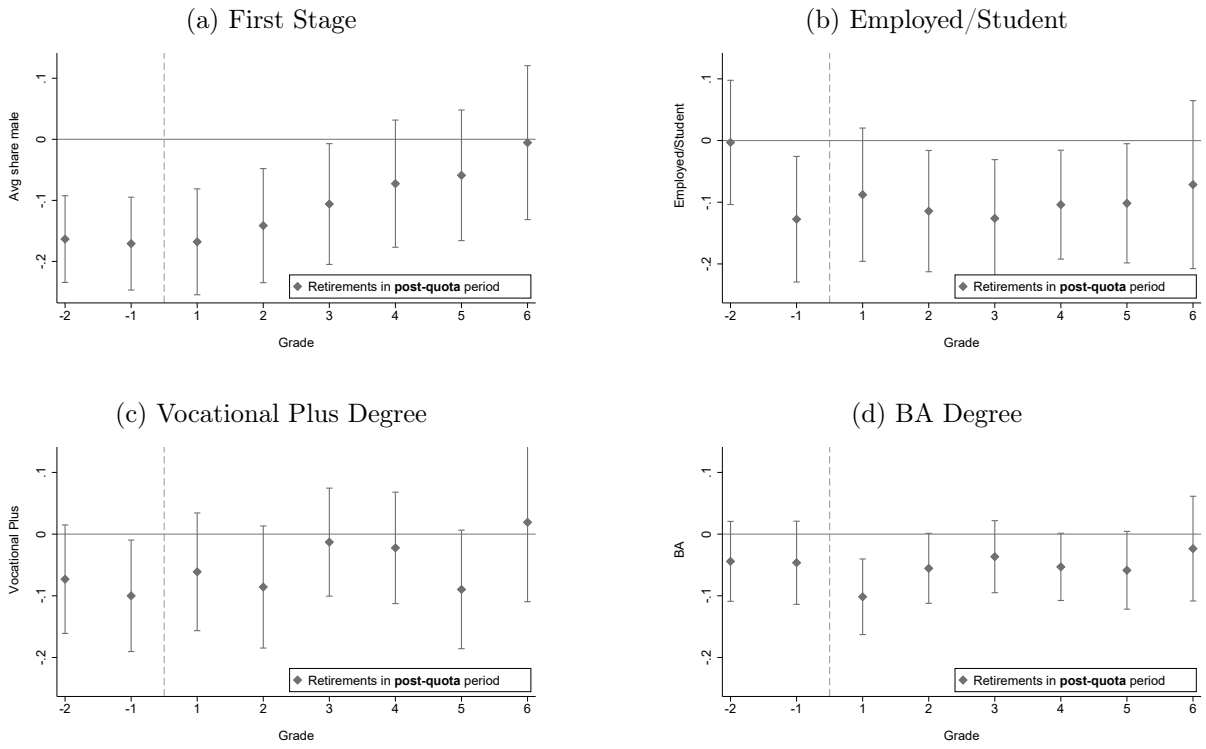
F.2 Reduced Form: Grade-Level for Selected Outcomes

Figure A6: Intermediate Outcomes: Grade Level Estimation



Note: Grade level estimation of pupil level first stage and reduced form (see Equation 10). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Figure A7: Long-Term Outcomes: Grade Level Estimation



Note: Grade level estimation of pupil level reduced form (see Equation 10). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

G Mechanism

G.1 Complementarities

Table A17: Reduced Form: Complementarities in Production

	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Low: Total Share	–0.086	0.060	0.026	–0.085	0.017	0.023	0.039
60 * Post	(0.038)	(0.034)	(0.016)	(0.049)	(0.038)	(0.017)	(0.028)
High: Total	–0.066	0.058	0.009	–0.080	–0.000	0.025	0.055
Share 60 * Post	(0.036)	(0.033)	(0.017)	(0.045)	(0.036)	(0.016)	(0.024)
Low: Total Share	0.008	–0.010	0.002	0.053	–0.006	–0.008	–0.033
60	(0.033)	(0.028)	(0.014)	(0.043)	(0.033)	(0.014)	(0.022)
High: Total	0.039	–0.033	–0.006	0.028	0.008	–0.008	–0.031
Share 60	(0.027)	(0.024)	(0.012)	(0.040)	(0.031)	(0.012)	(0.021)
Obs	819,498	819,498	819,498	805,895	805,895	805,895	805,895
Low: Dep mean	.904	.07	.026	.846	.082	.017	.053
High: Dep mean	.928	.057	.015	.832	.097	.018	.052
P-value	0.711	0.957	0.472	0.941	0.744	0.906	0.668

Note: Reduced form estimates for Equation 10 by median initial share male teachers in a municipality in 1990. P-values reported for test of equality of coefficients in the post-quota period. Outcomes are pupils' application choices and labor market status (see Tables 3 and 5). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

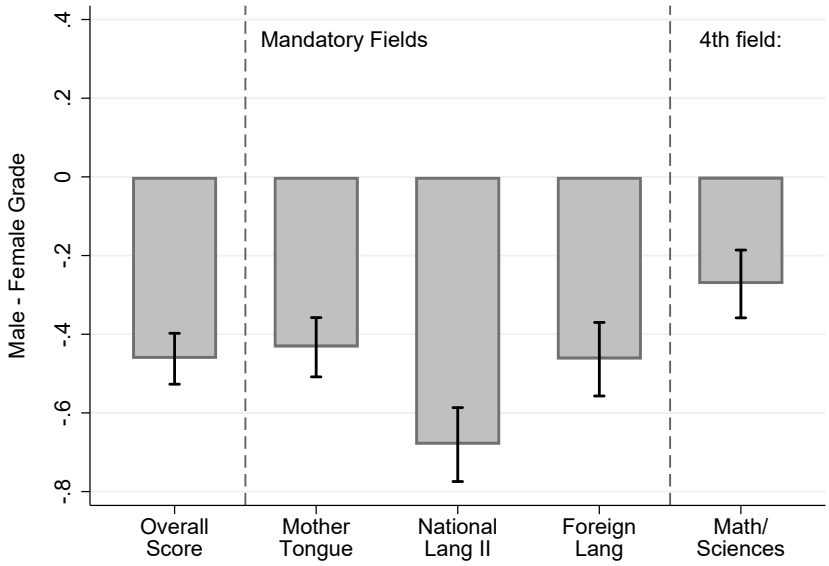
Table A18: Complementarities: IV estimates

	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Low: Avg Share	0.763	–0.536	–0.227	0.749	–0.149	–0.198	–0.344
Male	(0.578)	(0.456)	(0.191)	(0.593)	(0.340)	(0.195)	(0.292)
High: Avg Share	0.542	–0.472	–0.070	0.661	0.002	–0.208	–0.451
Male	(0.382)	(0.349)	(0.139)	(0.487)	(0.300)	(0.160)	(0.279)
Low: MOP F^{eff}	3.75	3.75	3.75	3.84	3.84	3.84	3.84
High: MOP F^{eff}	5.03	5.03	5.03	4.99	4.99	4.99	4.99
Low: Obs	590,156	590,156	590,156	579,101	579,101	579,101	579,101
High: Obs	229,342	229,342	229,342	226,794	226,794	226,794	226,794
Low: Dep mean	.904	.07	.026	.846	.082	.017	.053
High: Dep mean	.928	.057	.015	.832	.097	.018	.052
P-value	.75	.911	.506	.909	.74	.97	.791

Note: IV estimate of Equation 9, split sample by median share male teachers in a municipality in 1990. Outcomes are pupils' application choices and labor market status (see Tables 3 and 5). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

G.2 Performance Differences in Matriculation Exam

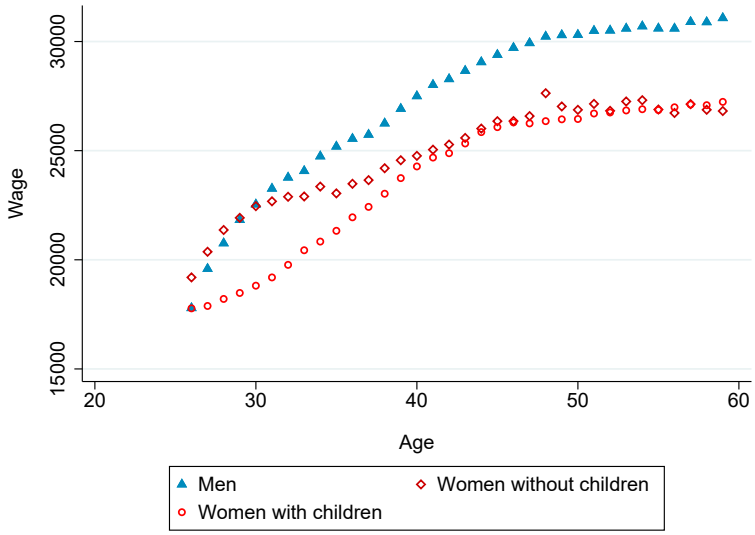
Figure A8: Performance Differences by Gender on Matriculation Exam by Field, Teachers



Note: Performance differences in the matriculation exam by gender for the full population of exam takers. Figure displays the gender difference in exam score ("Overall Score") and in each of the exam fields in the matriculation exam that count towards the overall score. Each displayed field receives equal weight of 25% in the calculation of overall score. The score for Math/Sciences is based on the best grade received if an exam taker chose both fields. Sample based on the full population of primary teachers in the main analysis sample who are first time matriculation exam takers in the high school track (lukion opiskelija) under age 22 for the years 1983 - 1985 (see Figure 8). Grades are assigned on a curve and range from 2-6, with 6 being the highest grade and 0 assigned for a failing exam (Kupiainen et al., 2018).

G.3 Career Attachment and Teachers' Compensation

Figure A9: Teacher Earnings by Age



Note: Wage earnings by age for the panel of active primary teachers. Earnings are winsorized below the 1st and above the 99th percentile.

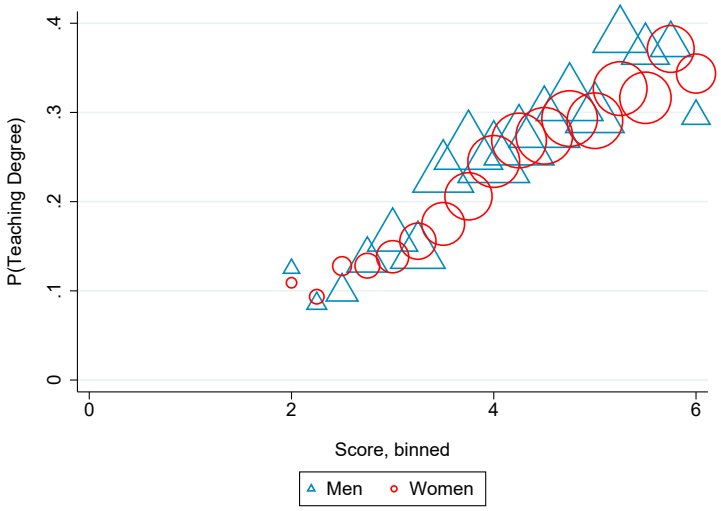
Table A19: Teacher Earnings

			No transfer	No Kids	Principal	
Male	3,164.808	3,021.450	2,592.573	1,046.137	0.036	2,906.805
	(26.840)	(44.990)	(25.982)	(84.721)	(0.002)	(67.183)
Experience		307.028				
		(1.445)				
Male * Experience		9.431				
		(2.497)				
Principal						7,920.115
						(225.695)
Adj. R^2	0.169	0.379	0.183	0.118	0.044	0.164
Obs	202,143	202,143	181,928	20,587	38,608	37,834
Dep mean	24,811	24,811	25,496	23,448	0.022	25,990
Dep mean men	26,980	26,980	27,269	24,134	0.046	28,178
Dep mean women	23,773	23,773	24,627	23,168	0.010	24,994

Note: Regressions based on the panel of active primary teachers (ages 26-60). Outcome in columns 1–4 and 6 are yearly earnings, outcome in column 5 is an indicator for being principal. No transfer: Sample based on active teachers who did not receive any government transfer in a given year. No kids: Sample based on teachers who never had a child. Experience is measured as time since degree. Earnings are winsorized below the 1st and above the 99th percentile. All specifications control for year and municipality fixed effects.

G.4 Compensating Differentials

Figure A10: Likelihood of Obtaining Primary Teaching Degree by Exam Score and Gender



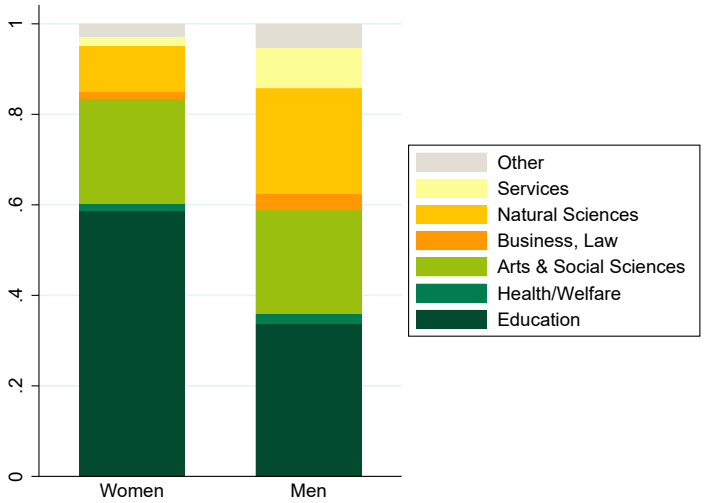
Note: Share of primary school teaching degree holders among all applicants to primary teaching by binned score in the matriculation exam. Weighted by number of applicants per bin. Sample contains all applicants to primary teacher education in the years 2000 - 2005, with data for 2003 missing and considers all degrees obtained by the year 2020. The exam score is based on a candidate's best performance if they had multiple takes. Grades are assigned on a curve and range from 2-6, with 6 being the highest grade and 0 assigned for a failing exam (Kupiainen et al., 2018).

Table A20: Likelihood of Teaching Degree by Score and Gender, Among Applicants to Primary Teaching

Score	0.072	0.073	0.069	0.070
	(0.003)	(0.004)	(0.004)	(0.004)
Male		–0.001		–0.006
		(0.036)		(0.043)
Male * Score		0.007		0.012
		(0.008)		(0.010)
Adj. R^2	0.100	0.100	0.082	0.084
Obs	20,560	20,560	17,934	17,934
Dep mean	.266	.266	.305	.305

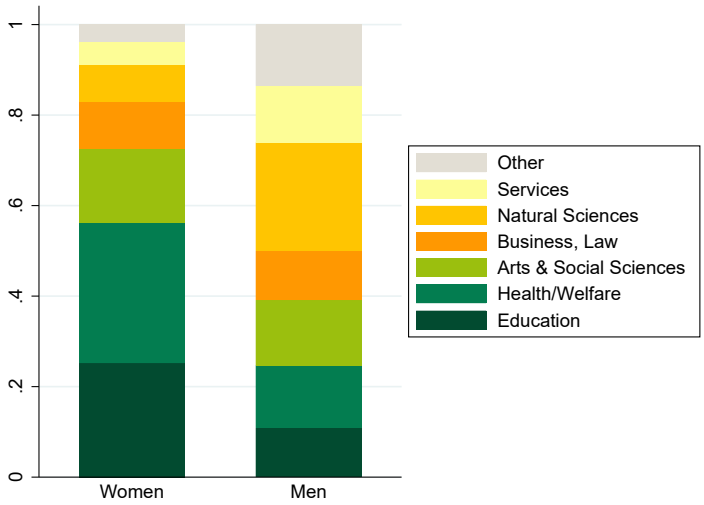
Note: Outcome is an indicator variable for whether an applicant holds a primary teaching degree in the year 2020. Sample contains all applicants to primary teacher education in the years 2000 - 2005 (data for 2003 missing) and considers all degrees obtained by the year 2020. In columns 3 and 4, sample is restricted to applicants who ever obtain any university degree. Controls for age, experience (since degree), and application year. If an applicant applies multiple years in a row during this time period, only the last application is considered. Exam score based on best performance if applicant had multiple takes. Score is assigned on a scale between 2-6 with 6 corresponding to the highest grade. Higher observation count relative to Table 9 due to considering degrees for each year up to 2020, whereas wages are measured in 2020 only (and thus some applicants do not reside in the country in that year).

Figure A11: Share of Fields Applied to Outside Primary Teacher



Note: Share of applications to fields other than primary school teaching among all applicants with at least one application to primary teaching. Sample contains all applicants to primary teacher education in the years 2000 - 2005 (data for 2003 missing).

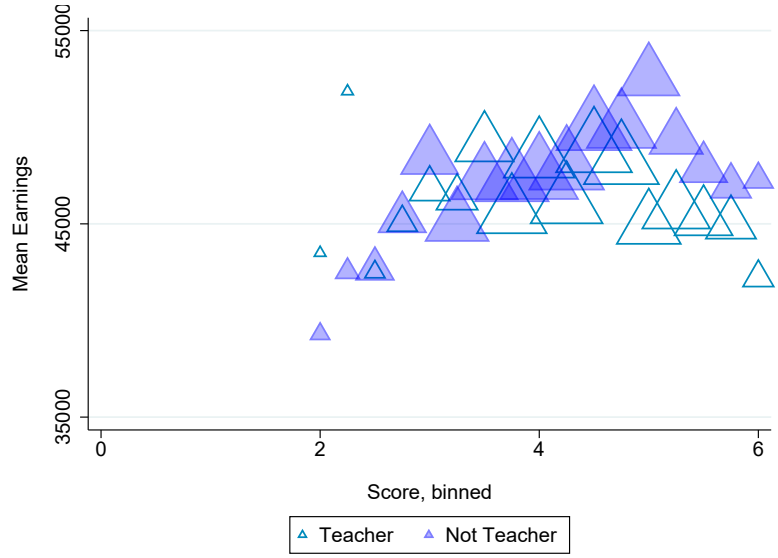
Figure A12: Actual Fields Obtained among Applicants Outside Primary Teaching



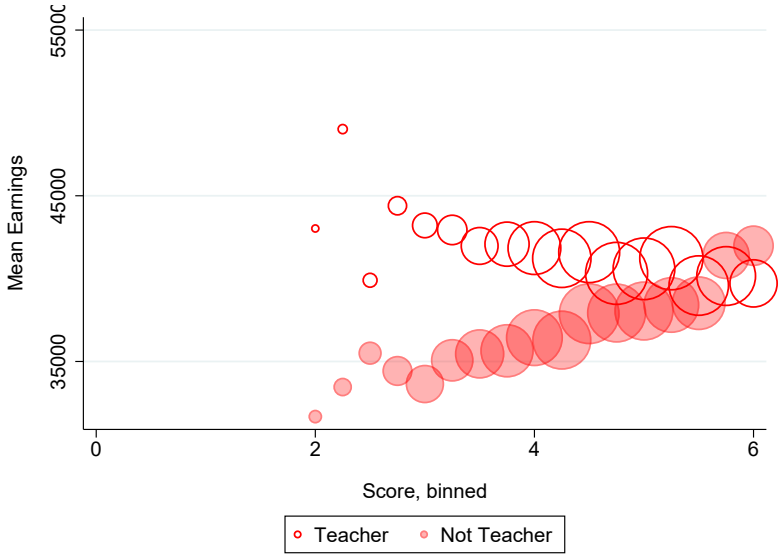
Note: Share of actual fields obtained among those who do not obtain a primary teaching degree by 2020. Sample contains all applicants to primary teacher education in the years 2000 - 2005 (data for 2003 missing) and considers most recent degree obtained in 2020.

Figure A13: Outside Options of Teacher Applicants: Residualized Earnings

(a) Men: Residualized Earnings



(b) Women: Residualized Earnings



Note: Residualized wage earnings in the year 2020 among all applicants to primary teaching by whether candidate holds a primary teaching degree against binned score in the matriculation exam. Weighted by number of applicants per bin. Controls for age, experience (since degree), and application year. Sample contains all applicants to primary teacher education in the years 2000 - 2005 (data for 2003 missing) and considers all degrees obtained by the year 2020. Exam score is based on best performance if multiple takes. Exam score is assigned on a scale between 2-6 with 6 corresponding to the highest grade.

G.5 Pupil Grades

Table A21: IV Estimates: Pupil Grades in Middle School and Matriculation Exam

	Middle School					Matriculation Exam				
	Overall GPA	Overall GPA < 99	Language	Math	Sample	Overall GPA	Language	Math	Math/ Science	Sample
Avg Share Male	–0.613	–0.648	–0.727	–0.006	0.038	–0.954	–0.671	0.659	–1.579	0.098
	(0.638)	(0.638)	(0.856)	(0.944)	(0.117)	(0.851)	(0.920)	(2.023)	(1.344)	(0.279)
MOP F^{eff}	15.78	19.87	19.92	19.91	15.35	16.11	16.11	17.01	16.12	15.44
Obs	800,811	674,913	670,046	669,352	825,095	395,674	395,674	318,288	395,706	811,393
Dep mean	7.616	7.608	7.548	7.485	.971	4.283	4.239	4.018	4.399	0.488
Std effect	-.035	-.037	-.038	0	.014	-.06	-.04	.026	-.076	.013

Note: Estimates of Equation 9. Outcomes from left to right: **Middle School:** Overall GPA across all subjects; Overall GPA for the cohorts that start primary school before 1999; Language: average across mother tongue, second national and foreign language; and Math. Language and Math are only available for the cohorts that start primary school before 1999. All grade information is based on pupils who ever apply to post-compulsory education and have a grade entered. Sample: Indicator for whether the pupil ever applies in the post-compulsory education data and has Overall GPA across all subjects entered. **Matriculation Exam:** Overall GPA across mandatory subjects; Language: average across mother tongue, second national and foreign language; Math; Math/Science: the better score across Math and Sciences. All grades are for pupils who pass the matriculation exam by age 20. Sample: Indicator for whether pupil is in Overall GPA sample for matriculation exam, excluding those who have died by age 25. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A22: Pupil Grades and Teacher Exam Scores

	Middle School				Matriculation Exam			
	Overall GPA				Overall GPA			
Overall GPA	–0.0075				0.0289			
	(0.0275)				(0.0405)			
Language		0.0042	0.0091			0.0815	0.0793	
		(0.0269)	(0.0284)			(0.0382)	(0.0395)	
Math or Science		–0.0139				–0.0619		
		(0.0216)				(0.0313)		
Math			–0.0035				–0.0180	
			(0.0120)				(0.0161)	
Science			–0.0167				–0.0291	
			(0.0198)				(0.0283)	
Math Background				–0.0225				0.0320
				(0.0607)				(0.0795)
Adj. R^2	0.188	0.188	0.188	0.188	0.081	0.081	0.081	0.081
Obs	800,750	800,750	800,750	800,750	395,643	395,643	395,643	395,643
Dep mean	7.616	7.616	7.616	7.616	4.283	4.283	4.283	4.283
Std effect X1	-.001	.001	.002	-.002	.005	.016	.015	.003
Std effect X2		-.003	-.001			-.013	-.006	
Std effect X3			-.004				-.007	

Note: Estimates of pupil outcomes on teachers' scores in the matriculation exam. Outcomes are pupils' Overall GPA across all subjects in middle school and Overall GPA in the matriculation exam (see Appendix Table A21 for a description of these outcomes). The score for Language is comprised of the (teachers') grade for Mother Tongue, Second National Language and Foreign Language. Math or Science is based on the best grade received if an exam taker chose both fields. Math background measures the share of teachers who have taken mathematics in their matriculation exam. Std effect X1-X3 indicate the standardized effect for each independent variable from (1) to bottom (3) for the respective column (i.e. in column 2, Std effect X1 indicates the standardized effect for Language score). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

H Robustness

H.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. Appendix Table A23 reports municipal level regressions, with all specifications assessing changes in flows for consistency (see Equation 7). Our goal is to understand whether teacher retirements in the post-quota period differentially impact municipalities' teacher hiring or exit strategies.

We start by assessing the effect of teachers turning 60 on the share of teachers leaving their current job in columns 1 and 2. Teachers turning 60 has almost a 1:1 impact on the share of teachers leaving, but not differentially so in the quota period. This effect is not driven by turnover of relatively younger teachers (column 2), and rather reinforces the observation that teachers reaching age 60 corresponds to actual exits from the teaching profession. In column 3 and 4, we examine how retirements affect proxies of experience of the local teacher body and do not detect a sizeable or significant change in the post-quota period, i.e. municipalities do not react to the policy reform by trying to recruit more teachers laterally. Column 5 shows (noisily estimated) that retirements in 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 teacher exit or re-hiring strategies as a response to the quota reform.

Table A23: Exit and Hiring Patterns in Municipalities

	Leave	Leave ≤ 59	Δ Age	Δ Time since degree	First entrants
Share 60 * Post-Quota	0.040	–0.003	–1.312	–1.245	–0.026
	(0.067)	(0.070)	(1.388)	(1.342)	(0.374)
Share 60	0.873	0.046	–16.090	–15.364	0.313
	(0.061)	(0.063)	(1.216)	(1.173)	(0.315)
Adj. R^2	0.176	0.008	0.222	0.206	0.037
Obs	4,448	4,448	4,448	4,448	3,745
Dep mean	.1	.08	-.21	-.21	.35

Note: Estimates for Equation 7. Outcomes from left to right: 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). All specifications include year fixed effects. Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted.

H.2 Teachers on Parental Leave

Apart from hiring patterns, the lifting the quota coincides with bringing more young female teachers to schools, who may have a higher propensity to go on leave when giving birth. The positive effects we detect from having more male quota teachers could then simply arise from pupils having less teacher turnover. During the 1990s, Finland provided 6.5 months of entirely shareable parental leave taking effect after three months of birth-related maternity leave (Kamerman and Moss, 2009). To check whether any changes related to leave taking of teachers becoming mothers (or fathers) could affect pupils, we repeat the municipal first stage regressions. Table A24 shows that teachers turning 60 in the post period do not have a differential impact on either female or male teachers having a birth in their household. The share of female teachers leaving the teacher force subsequent to becoming a mother is also not differentially affected by retirements in the post-quota period (column 4). In these specifications, the variation used stems from such patterns arising immediately as a response to teacher retirements. We therefore also document that, conditional on municipal and region-by-cohort fixed effects, higher exposure to female teachers having a newborn child does not impact pupil outcomes (Appendix Table A25). We conclude that leave taking patterns due to maternity from more female teachers post-quota are unlikely to drive our results.

Table A24: Teachers Becoming Parents

	Birth total	Birth fem	Birth male	Maternity leave
Share 60 * Post-Quota	0.064	0.033	0.031	–0.039
	(0.050)	(0.036)	(0.029)	(0.032)
Share 60	–0.121	–0.107	–0.014	–0.002
	(0.040)	(0.029)	(0.025)	(0.030)
Adj. R^2	0.006	0.006	0.012	0.005
Obs	4,448	4,448	4,448	4,448
Dep mean	.06	.03	.02	.02

Note: Estimates for Equation 7. Outcomes from left to right: Share of teachers with the birth of a child, share of teachers who are female and have a birth (column 2), and who are male and have a birth (column 3), share of teachers who are female and on leave after birth (defined as not being an active teacher in the year subsequent to having given birth). All specifications include year fixed effects. Standard errors clustered at the municipality level. Regressions weighted by population, means unweighted.

Table A25: Effect on Main Outcomes of Female Teachers Having a Newborn Child

	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Female teachers having a child	–0.009 (0.008)	0.007 (0.008)	0.002 (0.003)	–0.003 (0.010)	–0.007 (0.008)	0.006 (0.003)	0.003 (0.005)
Adj. R^2	0.070	0.030	0.138	0.025	0.023	0.003	0.013
Obs	825,095	825,095	825,095	811,393	811,393	811,393	811,393
Dep mean	.911	.066	.023	.842	.086	.017	.053

Note: Estimates of pupil outcomes on total exposure to female teachers having a newborn child while pupils are in primary school. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A26: Effect on Main Outcomes of Male Teachers Having a Newborn Child

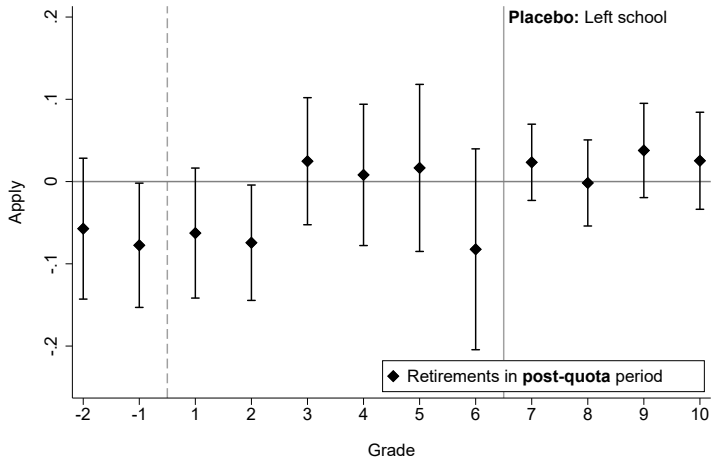
	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Male teachers having a child	–0.008 (0.010)	0.006 (0.009)	0.002 (0.005)	–0.021 (0.013)	0.009 (0.010)	–0.007 (0.004)	0.018 (0.007)
Adj. R^2	0.070	0.030	0.138	0.025	0.023	0.003	0.013
Obs	825,095	825,095	825,095	811,393	811,393	811,393	811,393
Dep mean	.911	.066	.023	.842	.086	.017	.053

Note: Estimates of pupil outcomes on total exposure to male teachers having a newborn child while pupils are in primary school. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

H.3 Placebo Test and Randomization Inference

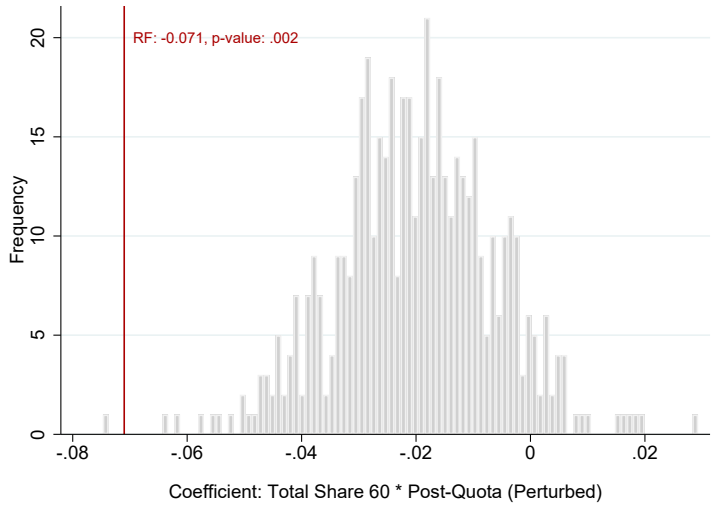
To further assess robustness, we perform a placebo test in Appendix Figure A14, by estimating reduced form treatment effects for additional "fake grades", i.e. years in which pupil cohorts have already left primary school. We show that any teacher retirements in those "fake grades" do not affect pupils' application outcomes. In addition, Appendix Figure A15 plots the distribution of reduced form treatment coefficients for applications (Equation 10), estimated from randomly re-assigning the share of retiring teachers across municipality*year cells 500 times for the post-quota period. This allows to assess significance without making parametric assumptions on the structure of the error term and renders a p-value of 0.002.

Figure A14: Placebo Reduced Form by Grade: Applications to Post-Compulsory Education



Note: Grade level estimation of pupil level reduced form (see Equation 10). Outcome is binary indicator for pupils applying to post-compulsory education directly after middle school, regressed on the share of teachers turning 60 just before a pupil enters the respective grade, including "fake grades" in which the pupil has already left primary school. Figure depicts coefficients for the post-quota period *relative* to the quota period. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Figure A15: Reduced Form: Randomization Inference



Note: Perturbed reduced form. 500 iterations of randomly assigning the share of teachers turning 60 in any post-quota year across municipalities, keeping fixed quota-period retirements, and estimating Equation 10. Outcome is binary indicator for pupils applying to post-compulsory education directly after middle school. Line indicates the reduced form coefficient in Table A14. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

H.4 Further Robustness

In Appendix Table Appendix Table A27, we show that attrition from our main sample is small and not selective by exposure to the quota reform. In Appendix Table A28 we document that

results are not driven by the capital (column 2) or large cities in general (column 3). Results are also not sensitive to defining treatment based on the actual municipality of living in each school year rather than our preferred approach which defines treatment based on the municipality of living at age 7. We further assess whether the impacts of male (quota) teachers are plausibly linearly additive in Appendix Table A29. In small municipalities, the municipal and school level share of male teachers coincide. We show that treatment impacts in small and larger municipalities are the same.

During the period of our study, two major macro-economic shocks warrant a brief mention: The depression in Finland during the early 1990s, as well as the financial crisis in 2008/09, initiating the global great recession. Recall that we study the cohorts born between 1981 – 1993, starting school between 1988 – 2000. Finland experienced a 14% contraction of GDP from 1990-93, accompanied by a more permanent rise in unemployment. Region-by-cohort fixed effects in all of our specification allow for differential impacts of this shock across different parts of the country. In addition, we run our main specification controlling for parental unemployment status at age 7 in columns 4 and 8 of Appendix Table A28, with the main effects quantitatively unchanged. Regarding the great recession, it is worth noting that our treatment assignment is based on the place where pupils live when they are age 7. Our study cohorts turn 25 years old in the years 2006-2018. It is thus the earlier and middle cohorts that are both more exposed to male quota teachers and turn 25 during the time of the financial crisis and subsequent recession.

Table A27: Test for Selective Sample Attrition

	RF		IV	
	Left 16	Left 25	Left 16	Left 25
Avg Share Male			0.026	–0.019
			(0.029)	(0.051)
Total Share 60 *	–0.004	0.003		
Post	(0.005)	(0.009)		
Total Share 60	0.006	0.006	0.005	0.007
	(0.005)	(0.007)	(0.004)	(0.006)
MOP F^{eff}			15.35	15.35
Adj. R^2	0.042	0.049		
Obs	826,515	826,515	826,181	826,181
Dep mean	.005	.018	.005	.018

Note: Reduced form, and IV estimate of Equation 9. Outcomes are a binary indicator for pupils having left the sample at age 16 or age 25, excluding children with registered deaths until age 25. Pupils are defined as having left the sample if they do not appear in the register data at the respective age. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A28: Further Robustness

	Apply					Employed/Student				
	Main	No capital	No cities	Parent UB	Movers	Main	No capital	No cities	Parent UB	Movers
Avg Share Male	0.418	0.431	0.452	0.388	0.428	0.510	0.489	0.501	0.457	0.492
	(0.195)	(0.196)	(0.212)	(0.191)	(0.206)	(0.242)	(0.236)	(0.254)	(0.233)	(0.240)
MOP F^{eff}	15.35	15.66	13.91	15.35	15.47	15.44	15.76	14.01	15.45	15.59
Obs	825,095	758,380	648,931	825,091	821,891	811,393	746,393	639,044	811,389	808,070
Dep mean	.911	.915	.918	.911	.911	.842	.841	.838	.842	.843

Note: IV estimate of Equation 9. Outcomes are pupils' applications in the last year of middle school (see Table 3) and labor market status (see Table 5), in turn examining the main specification for comparison (column 1 and 6), dropping Helsinki (column 2 and 7), dropping the five most populous municipalities based on place of living at age 7 (column 3 and 8), controlling for parental unemployment status at age 7 (column 4 and 9), and assigning treatment based on the municipality in which pupils live in each year during primary school instead of only the municipality of living at age 7 (columns 5 and 10). The lower observation count in columns 5 and 10 is due to excluding pupils for whom I cannot observe a municipality in all years. All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

Table A29: Reduced Form by Size of Municipality

	Apply directly	Apply late	Apply never	Employed/ Student	Un-employed	DI/ Pension	Other out of LF
Small: Total	–0.073	0.058	0.016	–0.083	–0.009	0.022	0.065
Share 60 * Post	(0.033)	(0.029)	(0.014)	(0.044)	(0.033)	(0.013)	(0.023)
Large: Total	–0.070	0.058	0.012	–0.087	0.010	0.021	0.053
Share 60 * Post	(0.028)	(0.025)	(0.012)	(0.033)	(0.026)	(0.011)	(0.019)
Total Share 60	0.017	–0.021	0.003	0.042	–0.000	–0.005	–0.035
	(0.021)	(0.019)	(0.009)	(0.029)	(0.023)	(0.009)	(0.015)
Adj. R^2	0.070	0.030	0.138	0.025	0.023	0.003	0.013
Obs	825,095	825,095	825,095	811,393	811,393	811,393	811,393
Small: Dep mean	.924	.055	.021	.849	.084	.016	.049
Large: Dep mean	.91	.067	.023	.842	.086	.017	.053
Share Small	.042	.042	.042	.042	.042	.042	.042
P-value	.892	.999	.727	.892	.429	.947	.507

Note: Reduced form estimates (Equation 10), estimating separate coefficients for small (12 teachers or fewer) and large municipalities (> 12 teachers). P-values for test of equality of coefficients in the post-quota period. Outcomes are pupils' applications in the last year of middle school (see Table 3) and labor market status (see Table 5). All specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH. Standard errors clustered at the municipality level.

H.5 TWFE Robustness

An active literature has documented that in the presence of heterogeneous treatment effects, the coefficient of a two-way fixed effects (TWFE) regression, $\hat{\beta}_{fe}$, may be a biased estimate of the treatment effect and in severe cases exhibit the opposite sign (See, among others: De Chaisemartin and d'Haultfoeuille, 2020; Arkhangelsky et al., 2021; Sun and Abraham, 2020; Imai and Kim, 2021; Goodman-Bacon, 2021). If treatment effects are heterogeneous, such bias arises when already treated units are used as a control group in later periods. In a two stage least squares (2SLS) set-up, potential issues would arise from residualized treatment assignment in the first stage (which is then used to generate predicted values of the endogenous variable for the second stage), *if* treatment effects are heterogeneous. In our setting, however, the first stage portrays a relationship between local retirements and teacher gender composition that should be purely mechanical, and for which — given our knowledge about the quota reform — we have a clear ex ante prior on sign and magnitude. While the TWFE literature to date has not tackled settings with continuous treatment variables, below we follow the reasoning outlined in De Chaisemartin and d'Haultfoeuille (2020) to discuss negative weights and potential heterogeneity in treatment effects. We further probe whether first stage coefficients are driven by particular years, regions, or levels of treatment assignment in Appendix H.6. We conclude that treatment effect heterogeneity leading to sign reversal in $\hat{\beta}_{fe}$ is not a major concern in our setting.

De Chaisemartin and d'Haultfoeuille (2020) decompose $\hat{\beta}_{fe}$ into a weighted sum of average treatment effects (ATE) for treated units, with weights proportional to the residual from a regression of the treatment variable on fixed effects. If treatment effects are heterogeneous, problems with sign reversal arise when treated observations receive a negative weight due to their *residualized* treatment value in a particular period being negative (intuitively, these negative weights arise because that particular observation serves as a control in that period). Negative weights by themselves are mechanically the product of any TWFE specification — it is in combination with heterogeneous treatment effects that problems with sign reversal may arise. We can use the intuition developed in De Chaisemartin and d'Haultfoeuille (2020) (also highlighted by Jakiela, 2021) to probe for such issues in our setting.

First, the highlighted concern arises only when treatment effects are heterogeneous. The way in which treatment effect heterogeneity matters in our IV set-up is through the first stage relationship by using residualized treatment assignment to generate predicted values for the endogenous variable in the second stage. The first stage in our setting estimates a mechanical relationship between local retirements and teacher gender composition, with a clear ex-ante prior on sign and magnitude. While there is no direct test of assessing treatment effect heterogeneity, reporting sensitivity to particular groups or time periods may at least be partly illuminating about whether the first stage coefficients are driven by any particular group of observations. To this extent, leave-one-out estimation in the following section (Appendix H.6) reports coefficients that show no worrisome patterns.

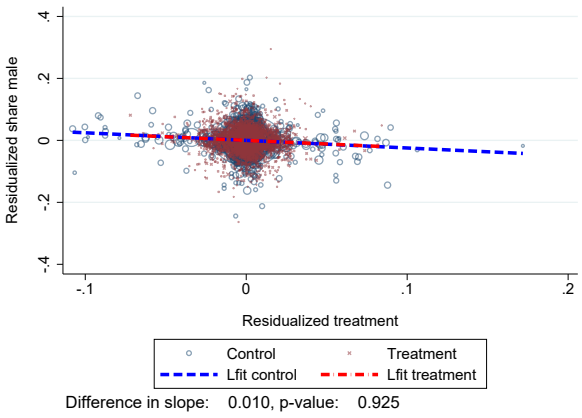
A further probing for treatment effect heterogeneity consists in examining the relationship between the residualized outcome and residualized treatment variable. Under homogeneous treatment effects, this relationship should be linear and not differ by treatment assignment status.

In the first stage of our setting, pupil cohorts that experience relatively more retirements in the post-quota period are 'treated' by being exposed to fewer male quota teachers, whereas pupils with relatively more retirements in the quota period serve as the 'control' group. Appendix Figures A16a and A16b plot the residualized share male against the residualized treatment variable both for the municipal and the pupil level first stage.[24] A test for differences in slopes between treatment and control observations shows that these are small and not significant.

Second, following De Chaisemartin and d'Haultfoeuille (2020), we can examine the weights that observations receive. The focus here is to understand how treatment assignment based on actual treatment status maps into treatment assignment based on the *residualized* treatment variable. In our setting, municipality-by-year or municipality-by-cohort observations with more retirements in the quota relative to the post-quota period should serve as a control group based on actual treatment assignment, and thus receive a negative weight (i.e. exhibit a negative *residualized* treatment assignment). Appendix Figure A17 plots residualized treatment assignment against actual treatment assignment both for the municipal and the pupil panel separately. Reassuringly, the mapping between residualized and actual treatment assignment follows a clear pattern: observations with higher retirement in the quota period are those that exhibit on average a negative residualized treatment value (i.e. receive a negative weight).

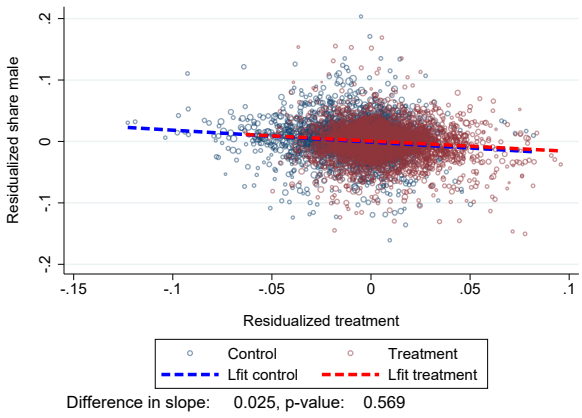
Figure A16: Treatment Effect Heterogeneity

(a) Municipality Level: Residuals of First Stage



Difference in slope: 0.010, p-value: 0.925

(b) Pupil Level: Residuals of First Stage



Difference in slope: 0.025, p-value: 0.569

Note: Residualized share male against residualized treatment (see Equation 15), at the municipal level (upper panel) and pupil level (lower panel). Weighted by number of observations.

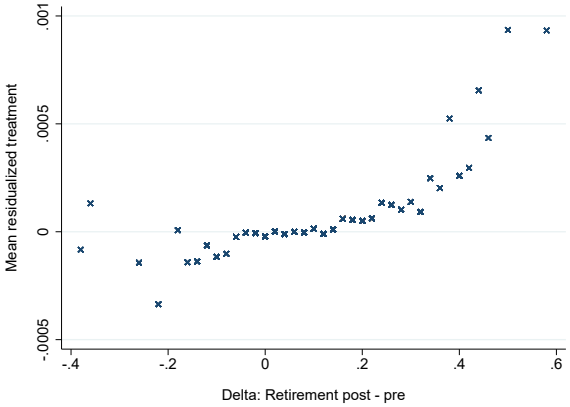
[24] I.e. for the first stage equation 8 at the municipal level, residualized treatment corresponds to the residuals of the following specification:

$$\textit{total share 60 post}_{mt} = \beta_0 + \beta_1 \textit{total share 60}_{mt} + \gamma_{mp} + \eta_t + \epsilon_{mt} \tag{15}$$

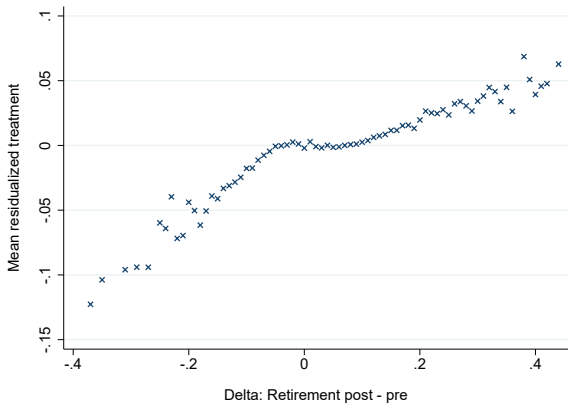
With $\textit{total share 60 post}_{mt}$ the share of teachers turning 60 interacted with an indicator for the post period.

Figure A17: Residualized Treatment

(a) Municipality Level: Residualized Treatment ("weights")



(b) Pupil Level: Residualized Treatment ("weights")

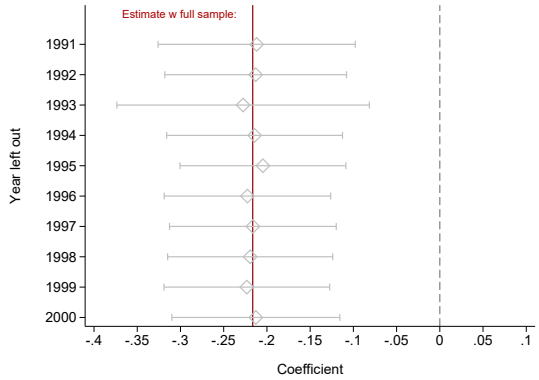


Note: Mean residualized treatment (see Equation 15) against actual treatment assignment (binned) at the municipal level (upper panel) and pupil level (lower panel).

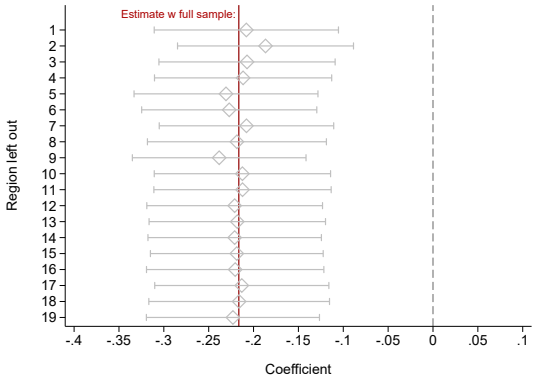
H.6 Leave-One-Out Estimation

Figure A18: First Stage: Leave-One-Out

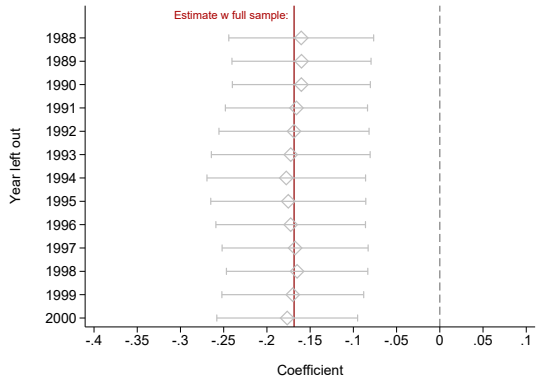
(a) Municipal Level: Leave-One-Out Years



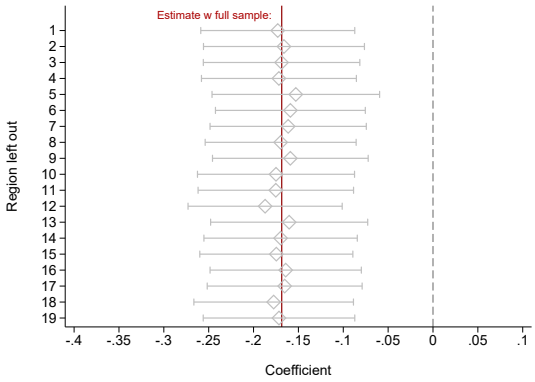
(b) Municipal Level: Leave-One-Out Region



(c) Pupil Level: Leave-One-Out Cohorts



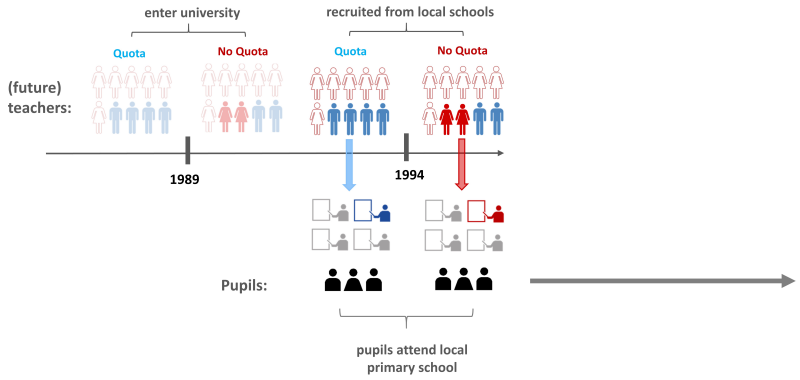
(d) Pupil Level: Leave-One-Out Region



Note: Leave-one-out estimation of treatment coefficient in municipal level first stage Equation 8 (Panel a and b) and pupil level first stage Equation 10 (Panel c and d), with respect to regions and years/cohorts. Indicated years/cohorts and regions on the y-axis are the respective observations dropped in the estimation of the coefficient. At the municipal level, all specifications include year and municipality-by-post-quota fixed effects. At the pupil level, all specifications include region-by-cohort and municipality fixed effects, as well as individual-level controls measured at age 7: gender, language (SE/FI/other), foreign origin, single parent HH, highest level of education in HH.

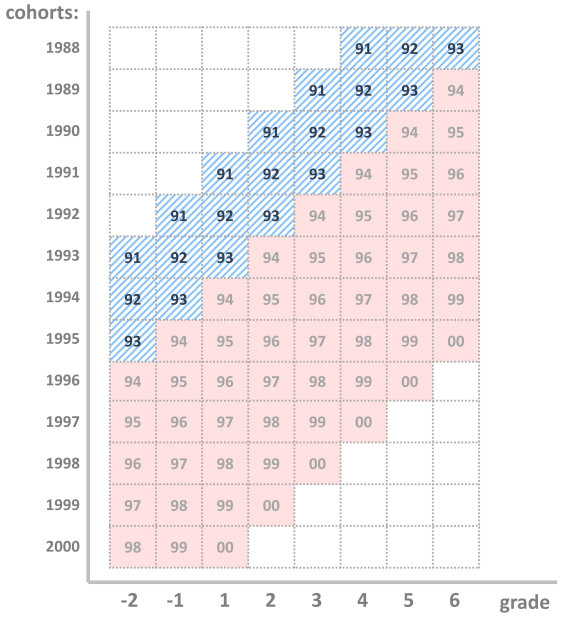
I Context

Figure A19: 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 before 1994 (quota), and thereafter (post-quota). 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 IV of the paper.

Figure A20: 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 (with stripes), years in which the quota was abolished in red.