NET WORK LISER

Discussion Paper Series

IZA DP No. 18604

April 2026

Dynamic Teacher Sorting

Rasmus H. Klokker
VIVE - The Danish Center for Social Science Research

Nicolai Kristensen
VIVE and IZA@LISER

IZA DP No. 18604 April 2026

Dynamic Teacher Sorting*

Abstract
Sorting among school teachers, whereby high-qualified school teachers often teach high-achieving students, is well documented, but relatively little is known about how this sorting arises. Based on Danish administrative records, we formulate and estimate a semi-structural dynamic model of career choice among 10 cohorts of teachers completing teacher college, whom we follow biannually up to 15 years. Among teachers initially in the public school, dynamic sorting results in high-qualified teachers moving towards teaching in high-achieving public schools or private schools, while the least-qualified teachers enter low-SES public schools and stay there. Teachers leaving the public or private school system are found to have certain time-constant unobserved characteristics that either immediately or over time make them choose other occupations.

JEL classification
I21, I24, I28

Keywords
teacher sorting, semi-structural estimation, dynamic model

Corresponding author
Nicolai Kristensen
nikr@vive.dk

* This paper was chapter 2 in Rasmus Klokker's PhD thesis from 2025.

Large gaps exist in our understanding of how teacher labor markets function.

Boyd, Lankford, Loeb & Wyckoff, *Journal of Labor Economics* (2013: 84).

1 Introduction

This paper is motivated by two important stylized facts about school teachers' labor market. Fact number one, teacher sorting, whereby high-achieving students are taught by the best-qualified teachers, is pervasive. Fact number two, many countries face a shortage of qualified public school teachers, OECD (2024).[1]

In this paper, we address both of these two stylized facts which characterize the teacher labor market in many countries. A large literature is devoted to teachers choice of occupation and preferences for school characteristics (e.g., Boyd et al., 2005; Jackson, 2009; Clotfelter et al., 2010; Hanushek & Rivkin, 2010; Boyd et al., 2013; Bonhomme et al. 2015) including why they may leave or never enter the teaching profession (e.g., Lindqvist & Nordanger, 2016; and the systematic review of 120 studies by Nguyen et al. 2020).[2] However, most of these studies are essentially static in nature even tough some studies do take a life-cycle or career-perspective or analyze teacher mobility (*i.a.* Hanushek et al., 2004; Podgursky et al. 2004; Boyd et al. 2008).

Further contributions to our understanding of the dynamics of teachers' occupational choices and teachers' sorting behavior over time include Wiswall (2013) and Goldhaber et al. (2023). Wiswall (2013) (re)-investigates teaching experience increases teacher quality, and finds the increases to be substantial and, for math teachers, higher than previously

[1]In the US, there are approximately 55,000 vacant positions and 270,000 teacher positions are filled with less qualified or uncertified candidates (see https://www.teachershortages.com/. In the UK, teacher vacancies are found for about 30 percent of schools in the most disadvantaged areas outside of London, and about 46 percent of schools in disadvantaged areas in London, Sibieta (2020).

[2]Most of the literature is based on data from the US. Non-US studies include a cross-country study by Luschei & Jeong (2018); and, Gensowski et al., (2024) using Danish data.

found. Goldhaber et al. (2023) use a simulation approach to find that if inequities in teacher mobility and hiring across different schools are removed, teacher quality gaps (TQGs) would close within 5 years. Eliminating inequities in teacher attrition without addressing mobility and hiring has much less impact on TQGs. This underscores the importance of teacher mobility and choice of occupation.

In this paper, we add to the literature that investigates dynamics of teacher sorting, in order to better understand the underlying mechanisms of teacher sorting and the choice-patterns teachers with different characteristics follow over time. To this end, we formulate a semi-structural model of occupational choices of teachers. We follow 10 cohorts of the population of graduates from teacher colleges in Denmark during the years 2008-2017, and model their occupational choices bi-annually over (up to) 15 years (29 semesters) after graduation, i.e. over the calendar years 2008-2022.

Given the shortage of qualified teachers in public schools, a shortage which has also been documented for Denmark (see AE Rådet, 2023; Kristensen, 2023), we give special attention to how dynamic sorting spells out in the public school system. The key employment opportunity for newly graduated teachers is the public school, and therefore we split this option into three levels defined by the average parental income at the school (lowest quarter, middle 50% and highest quarter). We show how the choice of socioeconomic ranking based on average parental income of public schools associates with alternative measures of socioeconomic status (SES).

We analyze the dynamics of teacher sorting by modeling a discrete choice multinomial model. In the model, work experience explicitly evolves within the model framework based on past choices, which may open (or close) future occupational possibilities, and where experience is occupational-specific. Teacher quality is measured by the grade point average (GPA) from teacher college and experience across different types of schools.

Close in spirit to the methodology of our study, Behrman et al. (2016) formulate a

structural dynamic model to analyze whether and how private voucher schools, commonly used in Chile, attract high qualified teachers who could otherwise work in a public school. They model the choice of whether to obtain a teacher degree as well as subsequent choices of working as a teacher or in outside options. Behrman et al. (2016) focus on the private-public school choice while we focus on college-educated teachers only and teacher sorting also *within* the public school system. Behrman et al. (2013) use four waves of survey data following 1,401 individuals. With access to the population of *all* graduates from teacher college over many years our data allow us to follow individuals in greater detail and over longer time spans.

Albeit methodologically different, Bruno et al. (2020) resembles our study as they use 10 years of administrative data from the Los Angeles Unified School District to describe patterns of newly educated teachers' sorting. They find that, relative to more experienced teachers, novice teachers have placements that are more challenging along several dimensions.[3]

A number of important findings emerge from our study. We find that college GPA matters greatly for occupational choices, equivalent to teacher sorting, but mostly in the first years upon graduation. In later years, experience may serve as a substitute for high GPA, and path the way to the most desired types of teaching jobs, i.e., top quarter of public schools or the private school option. Gradually over time, some of the best-qualified teachers leave the public school, and male teachers in particular leave the public school. A finding much in line with Wiswall (2013). Low SES public schools attract teachers with relatively low GPA and these teachers tend to stay in low SES public schools. Across time, some of the graduates we follow choose not to be a teacher, and this share rises from about 15% in early years to more than 20% after 15 years. A similar development is seen for the job option "other teaching" so that, jointly, 40% do not teach in neither the

[3]The dimensions included in Bruno et al. (2020) are intensity of instructional responsibilities, homophily, colleaguesâ qualifications, and professional culture at the school.

public nor the private elementary school 15 years after graduation from teacher college.

We simultaneously estimate the impact of the choices on the wage earnings over time. Even though, on average, some of the best-qualified college graduates turn either to private schools or out of the elementary school, they do not earn a wage premium by making these choices. Instead, they choose to indirectly pay a premium in terms of decreased wage earnings in order to not be a public school teacher.

A key advantage of structural model estimation is that we can simulate effects of counterfactual policies. Here, we simulate the impact of a hypothetical policy whereby newly graduated teachers are submitted to a grace period of 2 years in which they have to teach in the public school system (maintaining an outside option of no employment as a reference group). The counterfactual simulation reveals that such a short-lived grace period may have permanent effects, but less so for the low SES public schools. And, the permanent effects also include a substantial permanent increase in teachers leaving the teacher profession altogether.

The remainder of the paper is organized as follows. In section 2 we describe the data. This is followed by a brief description of relevant theories in labor economics. In section 4 and 5 we describe the model and estimation approach. Sections 6 and 7 include a description of estimation and simulation results. In section 8 we discuss the implications of our findings and relate them to existing studies. Section 9 includes concluding remarks and suggestions for future research.

2 Data

2.1 Administrative employer-employee panel data

Our study is based on comprehensive, high-quality administrative records from Statistics Denmark based on the Central Personal Registry (CPR). Each individual has a unique

CPR number, and this number is matched to all administrative registries, which makes it possible to follow individuals over time and across different types of registries. Also, every workplace has its own firm and plant identifiers and can be linked with all its employees over time. In this paper, we connect individual panel data, where, from 2008 and onward, we observe job earnings and income transfers on a monthly basis, which we use to construct bi-annual data. We use August to measure the fall semester and February to measure the spring semester.[4]

The data includes detailed information on demographics, education, and labor market activities.

2.2 Measure of teacher quality

Goldhaber et al. (2023) note how the teacher quality gaps (TQGs), which arise when there is teacher sorting, are evident irrespective of whether teacher quality is measured by experience, degrees, or other advanced credentials (e.g., Bruno et al., 2020; Rodriguez et al., 2023) or by value added measures of teacher quality (e.g., Goldhaber et al., 2015; Isenberg et al., 2022; James & Wyckoff, 2022).

In this study, teacher quality is measured primarily by the GPA from teacher college. The GPA, like any other succinct measure, clearly has its limitations. Marra (2021: 14) discuss elements that go into quality teaching and categorize them into six characteristics: Knowledge, Passion, Compassion, Authority, Character and Organization. Clearly, these varying characteristics cannot be fully measured by college GPA nor by credentials. Still, the GPA from teacher college provides a measure of teaching skills such as classroom leadership ability, the ability to build teacher-student relations, as well as didactic skills. In other words, while teachers cognitive skills likely are positively associated with college GPA it is clear that college GPA provides a much broader and much better measure that

[4]We use the main occupation only.

also takes more "soft" skills into account. In addition to college GPA, teacher quality is also measured by occupational-specific experience accumulated over time.

2.3 The Sample

We follow the entire population of graduates from teacher colleges in Denmark over the years 2008-2017. This amounts to 24,941 new teachers. Subsequently, we restrict the sample. Teachers who permanently exit the labor market have been excluded from the sample. This includes individuals who receive early retirement pension. Table 2 shows how our sampling requirements affect the sample.[5] Teachers with alternative degrees, so-called "merit teachers" are excluded from the sample as are teachers for whom we do not observe their GPA from teacher college.

The final sample of teacher graduates includes 21,343, whom on average we observe for 21.8 periods. To reduce the computational burden we randomly select 10,000 teacher graduates out of the 21,343, leaving a a total of 217,700 observations. This is the sample we use in the analysis.

Since the sample is based on different cohorts of college graduates it is inherently unbalanced. However, we did require each person to be observed at a minimum of 5 years, so for the first 10 periods (semesters) all 10,000 teachers are observed, while only those from the very first cohort of 2008-graduates potentially are observed i all 29 periods. In this case this amounts to 945 individuals.

The reference group in our model will be teachers who choose the option called "Out". This option includes teacher graduates receiving unemployment insurance(UI), sick leave, or parental leave. The "Out" option also includes teachers who are not present in the data in a given period in the first 10 periods.[6]

[5]For our cohorts, post-employment wage was feasible at age 63.

[6]For instance, a teacher graduate may be observed in periods 1-4, be unobserved in period 5, and then subsequently be observed in periods 6-10. The teacher graduate would then be classified as "Out"

in period 5.

Table 2: Steps taken in data processing and their impact on sample size

Action taken	Sample size	Percent decrease in sample size
All graduates from teacher colleges 2008-2017	24,941	NA
Remove teachers with alternatively certificate	23,132	7.3%
Remove teachers unemployed across the entire period	23,097	0.2%
Remove teachers who have 'gaps' in the administrative data	23,091	0.0%
Restrict to teachers younger than 63 in 2022	22,781	1.3%
Restrict to teachers with observed college GPA	21,398	6.1%
Restrict to teachers observed in periods since graduation	21,343	0.3%

2.4 A first look at data

Table 3 shows how the 10,000 individuals are distributed across the 7 occupational groups that we model. Period 1 is the first half year following graduation. In other words, teachers graduating in June will be in period 1 the following August, period 2 the next years February, and so forth. The Periods cut across calendar time as we have stacked the data with respect to periods following graduation.

Since we are interested in teacher sorting, including sorting toward the most popular public schools, we allocate each public school into one of three "types" according to the average parental income at the school. "Public 0-25%" refers to schools where average parental income is in the lowest quartile of all public schools; "Public 25-75%" are the middle income schools, and "Public 75-100%" are schools where parental income is in the top quarter. In period 1, 57% enter one of the three public school categories. In period 15, the same share is 53% while in period 29 it has dropped to 48%. The share working in the high SES public schools is largely constant, while the graduates in our sample over time shy away from medium or low SES schools.

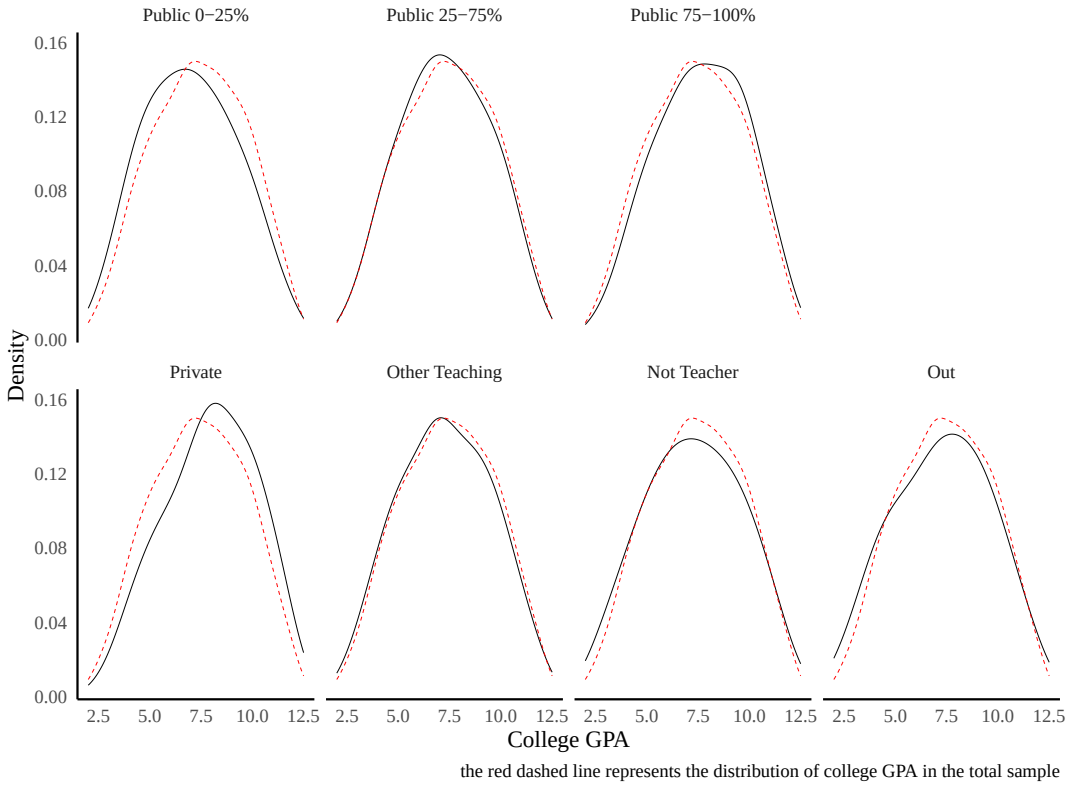
Private schools include boarding schools and so-called "free schools". These schools occupy 12-14% of the sample and this share is stable over the periods. So is the "out" option, which includes unemployed. It increases slightly in period 22-29 but it appears relatively stable. Instead, the teachers leaving elementary school, notably the public schools, either turn to "other teaching" or choose not to be a teacher. In period 1 these two options jointly account for 22% of the teachers while from period 20 the share has increased to roughly 30%.

Table 3: Percentage of sample in each status, by periods after graduation

Period	N	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out	Total
1	10,000	0.14	0.26	0.17	0.13	0.14	0.08	0.08	1
2	10,000	0.14	0.27	0.17	0.14	0.14	0.06	0.08	1
3	10,000	0.14	0.27	0.17	0.14	0.14	0.06	0.07	1
4	10,000	0.14	0.27	0.18	0.14	0.14	0.06	0.07	1
5	10,000	0.14	0.26	0.18	0.14	0.15	0.06	0.07	1
6	10,000	0.14	0.26	0.18	0.14	0.15	0.06	0.07	1
7	10,000	0.14	0.26	0.18	0.15	0.15	0.05	0.07	1
8	10,000	0.14	0.25	0.18	0.14	0.16	0.06	0.07	1
9	10,000	0.13	0.25	0.18	0.14	0.16	0.06	0.07	1
10	10,000	0.13	0.25	0.18	0.14	0.16	0.07	0.07	1
11	9,923	0.13	0.25	0.18	0.14	0.17	0.07	0.07	1
12	9,172	0.13	0.25	0.17	0.14	0.17	0.07	0.07	1
13	9,042	0.13	0.24	0.17	0.14	0.18	0.07	0.07	1
14	8,152	0.12	0.24	0.17	0.14	0.18	0.08	0.07	1
15	8,014	0.12	0.24	0.17	0.14	0.18	0.08	0.08	1
16	7,042	0.12	0.24	0.16	0.14	0.19	0.08	0.07	1
17	6,892	0.12	0.24	0.16	0.13	0.19	0.08	0.08	1
18	5,983	0.11	0.23	0.16	0.13	0.19	0.09	0.08	1
19	5,860	0.11	0.23	0.15	0.13	0.19	0.09	0.08	1
20	5,046	0.11	0.23	0.15	0.13	0.20	0.09	0.08	1
21	4,928	0.10	0.23	0.16	0.13	0.20	0.09	0.08	1
22	4,184	0.11	0.22	0.16	0.13	0.20	0.09	0.09	1
23	4,033	0.10	0.22	0.16	0.13	0.21	0.09	0.09	1
24	3,314	0.11	0.22	0.15	0.13	0.20	0.09	0.10	1
25	3,039	0.11	0.22	0.15	0.13	0.21	0.09	0.09	1
26	2,204	0.10	0.22	0.15	0.12	0.21	0.10	0.10	1
27	2,030	0.10	0.21	0.16	0.12	0.22	0.10	0.09	1
28	1,151	0.09	0.22	0.16	0.12	0.20	0.10	0.12	1
29	945	0.09	0.22	0.17	0.12	0.19	0.11	0.10	1

In Figure 1, the black lines in each of the panels signify the distribution for each particular occupation while the red dashed lines signify the overall average GPA. These are the same across each of the seven panels but serve to ease comparison between occupations. Teachers choosing the "public 0-25%"-option are below the average GPA and teachers choosing "public 75-100%" are above the average GPA as are private school teachers.[7]

Figure 1: The distributions of college GPA across occupations



Given our special interest in teacher sorting into the public schools, it is illustrative to compare the share of teachers above (below) the median college GPA over time. The left-most panel of Figure 2 illustrates how the share of teachers with a college GPA below the median is about 57% in period 1 in the "public 0-25%" option, and as the teachers experience increases with periods increasing, it is the teachers with GPA below the median

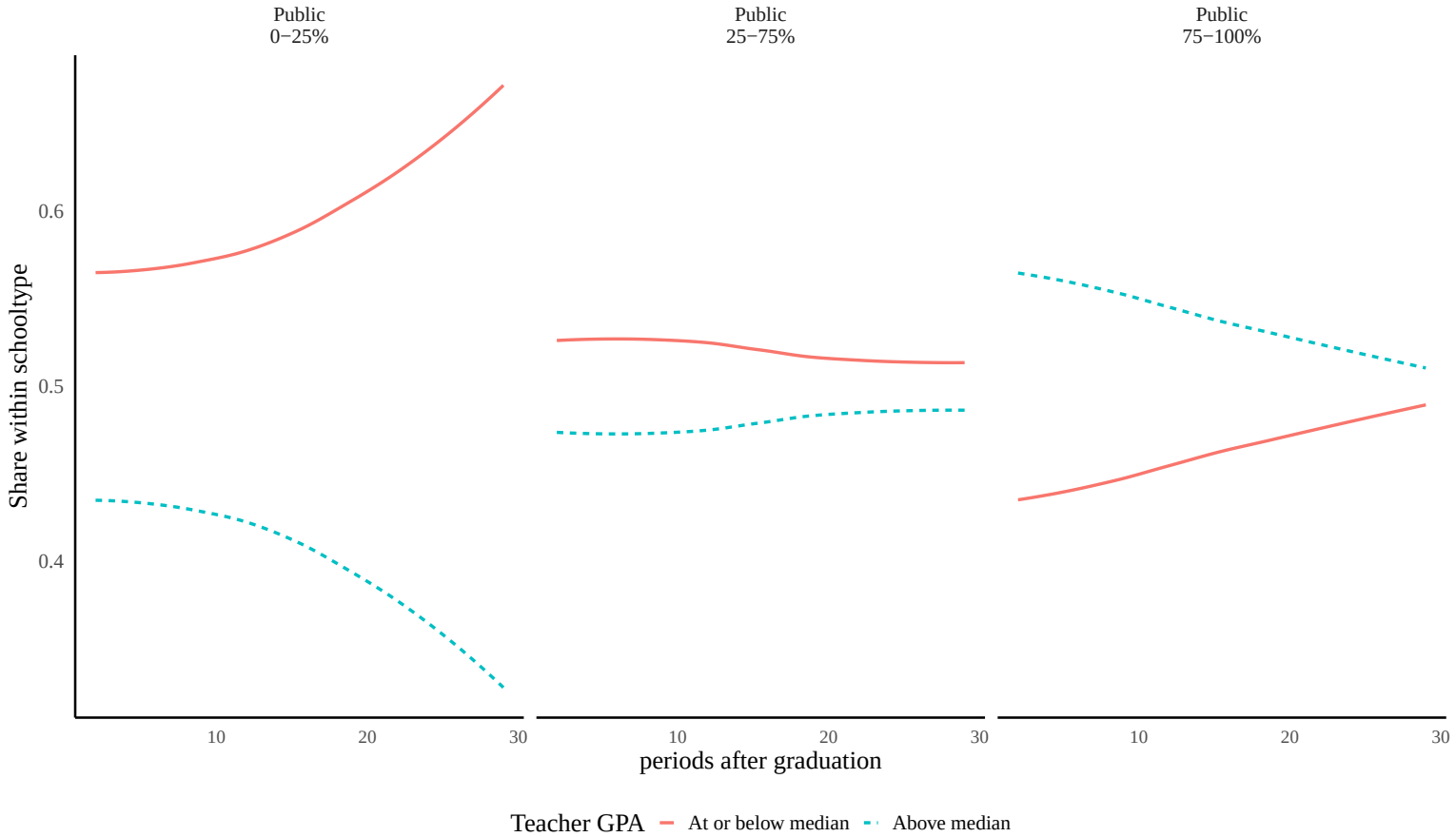
[7]Figure 1 is based on choices across all periods.

who remain in the low SES public schools (the "public 0-25%" option). The share with a GPA below the median increases by some 10 percentage-points for this school-type.

The mid-panel of Figure 2 shows the split between above (below) median college GPA in the middle 50% of schools as measured by school SES. For this occupational option there is almost no difference between the share above and below the median. This is to be expected. Also note how the difference narrows with increasing period length.

Finally, the right-most panel of Figure 2 shows how the high-SES schools are able to attract newly educated teachers who more often than not have a college GPA above the median. As experience increases, some of these high GPA-teachers leave the public school system and in the high SES occupational option (the "public 75-100%" option) they are replaced by teachers with below-median college GPA but who to some degree have experience *in lieu* of a high college GPA.

Figure 2: Share of teachers above (below) the median GPA, by public school SES

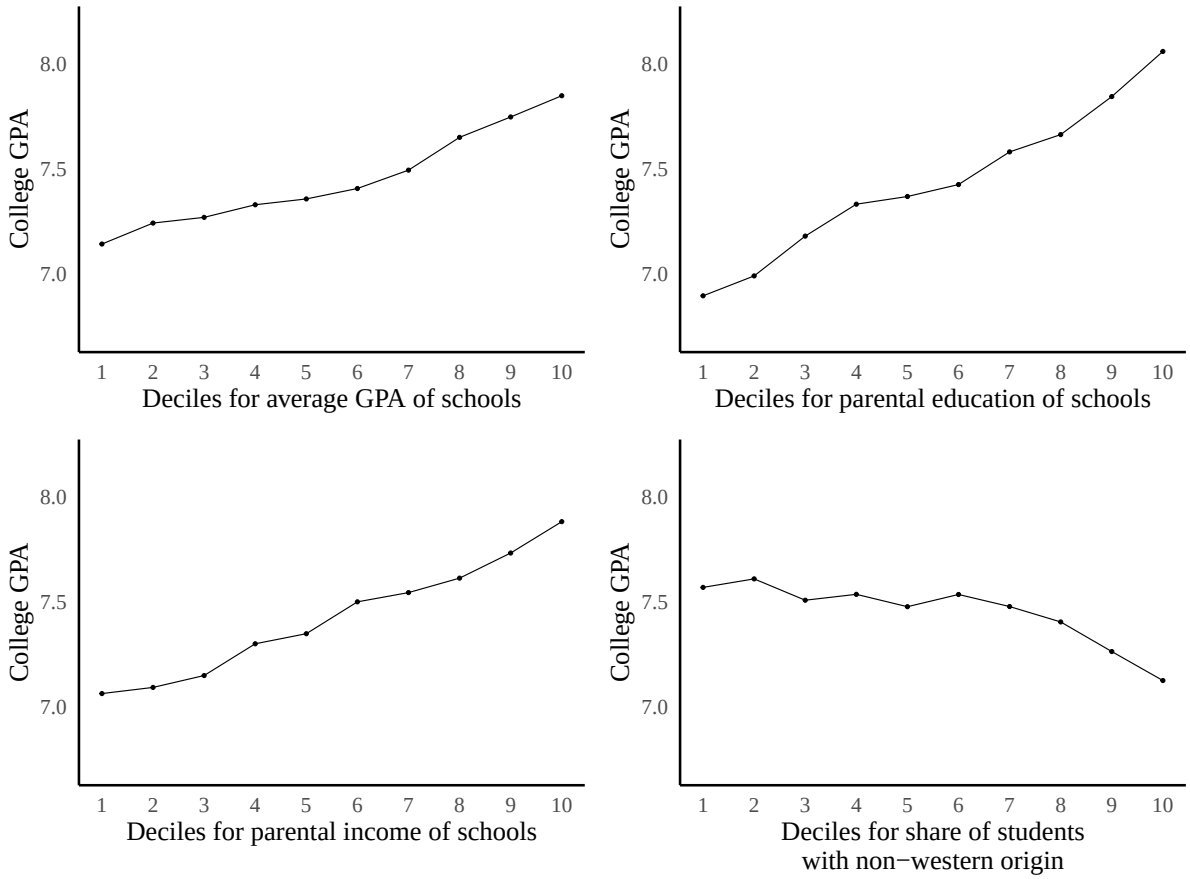


As mentioned above, the public schools are split by average parental income at the school level. This is only but one manner in which one can illustrate teacher sorting across school SES. We also considered defining school SES by i) the students GPA at final lower secondary exam at the school; ii) parents length of education; or iii) share of non-western immigrants. In Figure 3 we show how the teacher college GPA develops over these three alternatives and the chosen definition based on average parental income.[8] With the exception of non-western immigrants, the pattern of teacher sorting becomes clear in three of the four panels of Figure 3. The gradient is steepest for parental education but the difference between the three alternatives is relatively minor. This underscores how an occupational choice in a low SES public school ("public 0-25%") comes with a "package" of characteristics that differ from higher SES alternatives. One notable difference is that the share of non-western immigrants co-vary only very little with teacher college GPA's. From decile 1-7 the teacher college GPA is virtually constant at approximately 7.5 and it then decreases only slightly.[9]

[8]In the model, we do not split the public school option into school SES deciles as this would introduce a very large computational burden, but the deciles visually illustrate teacher sorting.

[9]Compared to, e.g., Hanushek et al. (2004) where race plays an important role, this lack of correlation is surprising.

Figure 3: Average College GPA across Deciles of School SES Characteristics



Note: *Top left*: The average GPA is computed as the school average in final exams in upper secondary school. For students from feeder schools, their subsequent average from the receiving upper secondary school is applied and traced back to the feeder school. *Top right*: Parental education is measured according to the nominated time for completion their highest education (called *pria*) measured in months. *Bottom left*: Parental income is based on the gross income of both parents (irrespective of cohabitation). *Bottom right*: The definition of non-western immigrants follow the definition by Statistics Denmark.

Given our dynamic modeling approach it is instructive to observe transitions between occupations from period (t-1) to period t. Looking at the average transitions in Figure 4, the expected strong time-lag emerges as very high shares in the diagonal of the transition matrix. The matrix also reveal how very small shares move from higher public SES downwards or from private schools to public schools. Only 1% of teachers from a high SES public school ("public 75-100%") in period (t-1) transit to low SES public schools

in period t, while 4% transit to a middle-SES public school. These modest shares of transition to low SES public schools are primarily observed during the first 10 periods after which teachers in high SES public schools virtually never move to a low SES public school (see the transition matrices by groups of periods included in the appendix).

Table 4: Transition matrix

	status in period (t)							
status in period (t-1)	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out	Total
Public 0-25%	0.87	0.06	0.01	0.01	0.03	0.00	0.02	1
Public 25-75%	0.03	0.89	0.03	0.01	0.03	0.01	0.02	1
Public 75-100%	0.01	0.04	0.90	0.01	0.02	0.01	0.02	1
Private	0.01	0.01	0.01	0.93	0.02	0.01	0.02	1
Other Teaching	0.02	0.03	0.02	0.01	0.87	0.02	0.03	1
Not Teacher	0.01	0.02	0.01	0.01	0.05	0.84	0.05	1
Out	0.02	0.05	0.03	0.03	0.07	0.06	0.73	1

Note: Transitions from t-1 to t in total over all periods.

3 Theory

Salop & Salop (1976) were among the first to introduce the idea of employees' sorting behavior, and the sorting concept features as an intrinsic theoretical component in the entire literature on Personnel Economics developed by Lazear (e.g., Lazear (1998, 2005), Lazear & Shaw (2007), Lazear & Oyer (2013)). According to the theory of sorting behavior, an individuals choice of job depends on the utility derived from a given job opportunity vis-à-vis the utility derived from all feasible alternatives. Job utility depends on a series of factors, such as remuneration, job security, job content, and working conditions. This means that we can assign non-pecuniary utility components a monetary value. Empirical support of this idea can be found in, e.g., Oyer & Schafer (2005), Eriksson & Kristensen (2014). The latter also finds empirical support of the theoretical result that working

conditions induce employee sorting behavior. Personnel Economics provides a theoretical foundation for sorting (also on job content), but in fact, the entire literature on search frictions (e.g., Mortensen & Pissarides, 1999) is also closely related to the notion of employee sorting.

4 Methodological approach

The model developed in this section shares similarities with Ashworth et al. (2021) and Hansen & Kristensen (2024). The model forms the basis for the econometric specification which is designed to account for dynamic selection into activities and the associated endogeneity of types of accumulated work experience in the estimation of the utility of various types of occupation as well as their wage returns.

4.1 Occupational choices

In each period, which is measured as a 6-month period, each individual i chooses occupation j from a set of J possible occupations. The occupations we model are mutually exclusive:

1. Teacher in low SES public school ($d_{it}^{ses1} = 1$)
2. Teacher in middle SES public school ($d_{it}^{ses2} = 1$)
3. Teacher in high SES public school ($d_{it}^{ses3} = 1$)
4. Teacher in private school (incl. boarding schools) ($d_{it}^{p} = 1$)
5. Other types of teaching, including teaching in high school or teaching adults ($d_{it}^{ot} = 1$)
6. Non-teaching job including self-employed ($d_{it}^{nt} = 1$)

7. Out, this is the reference group which includes unemployed ($d_{it}^{out} = 1$)

In practice, in order to ensure that teachers are allocated to one unique occupational activity each semester, i.e., that the activities are mutually exclusive, we use the main occupation in February and August as the choice of that semester.

4.2 Value functions

The value function for individual i from choosing activity j in period t is denoted V_{it}^{j}. The value functions observed individual characteristics ($\mathbf{X}_{it}$); (ii) local unemployment in period t by gender and age, $\mathbf{UE}_{it}$; (iii) previous period's choice ($\mathbf{D}_{\mathbf{i,t-1}}^{\mathbf{j}}$); (iv) occupation-specific accumulated experience ($\mathbf{AE}_{\mathbf{i,t}}^{\mathbf{j}}$); and (v) unobserved, individual characteristics (α_i^j). The characteristics included in X are gender, GPA from teacher college, indicator for post 2014 reform, and an indicator for part time. We follow Ashworth et al. (2021) and Bernal and Keane (2010) and approximate the value functions as linear functions of these variables.

The model is greatly simplified because initial conditions can be ignored. All individuals complete college and we only need to model choices made after graduation where we follow all individuals closely period-by-period. And, all value functions contain the same type of variables in all periods.

The value function for each occupational choice j can therefore succinctly be written in a general form as:

$$V_{it}^{j} = \beta_x^j \mathbf{X}_{it} + \beta_{ue}^j \mathbf{UE}_{it} + \beta_d^j \mathbf{D}_{it-1} + \beta_{ae}^j \mathbf{AE}_{it}^j + \alpha_i^s$$

where $j \ \epsilon \ (1..7)$ so that

$$\mathbf{D}_{i,t-1} = \left(d_{i,t-1}^{ses1}, d_{i,t-1}^{ses2}, d_{i,t-1}^{ses3}, d_{i,t-1}^{p}, d_{i,t-1}^{ot}, d_{i,t-1}^{nt}, d_{i,t-1}^{out}\right)$$

are indicators for choices in the previous period. In any period and for each individual, only one of these indicators is 1 (and all other are zero).

The accumulated occupational specific experience terms, ($\mathbf{AE^j_{i,t}}$), augment the experience in occupational choice $\mathbf{j}$ by 1 in each period.[10] Hence, accumulated experiences evolve over time according to the following law of motion:

$$E^j_{i,t} = AE^j_{i,t-1} + d^j_{it}$$

where all experiences are zero at the beginning of period 1.

We normalize the value function to zero for the reference group "Out".

4.3 Wages

Wages are modeled as a function of the choices made and the choice-specific work experience, local labor market conditions, observed individual characteristics and unobserved heterogeneity. The wage function ω_{it} is defined as:

$$\omega_{it} = \delta^\omega_x \mathbf{X}_{it} + \delta^\omega_l \mathbf{ue}_{it} + \delta^\omega_q \mathbf{D}^w_{i,t-1} + \delta^\omega_a \mathbf{AE}_{i,t} + \alpha^\omega_i$$

4.4 Unobserved heterogeneity

In the equations above, we include college GPA as the measure of teacher quality. In part, this also reflects both cognitive and non-cognitive skills (as described in the data section), but it does not completely represent these skills. In addition, personality characteristics such as motivation and inner-drive may matter greatly. We therefore introduce correlated unobserved time-invariant, discretely distributed, unobserved heterogeneity terms

[10] Ashworth et al. (2021) is one of few studies that allow for more than a single-dimensional experience.

to each of the value functions above as well as to the wage function (in the vein of Heckman & Singer, 1984) These are represented by the α parameters in the equations above ($\alpha_i^j, j = ses1, ses2, ses3, p, ot, nt, out$). We model the joint distribution of the unobserved heterogeneity terms as finite mixtures of sets of α parameters which are assumed to be orthogonal to the included observed characteristics in each equation. In other words, there are K types of individuals, and each type k is endowed with a unique vector ($\alpha_k^j, j = ses1, ses2, ses3, p, ot, nt, out$).

5 Estimation

These value functions essentially amount to period-specific multinomial logit specifications augmented with endognous development of experiences and jointly estimated with the wage function.

We estimate the model parameters using maximum likelihood and add idiosyncratic stochastic terms (ε) to each of the equations above. We adopt a common assumption in the literature that the ε's have a type I extreme value distribution. In each period, individual i is assumed to choose the activity that yields the highest value from all activities and we can write the probability for individual i of choosing option j ($j \in ses1, ses2, ses3, p, ot, nt, out$) in period t, conditional on unobserved heterogeneity (α) as:

$$Pr(d_{it}^j = 1|\alpha_k^j) = \frac{exp(V_{it}^j(k))}{\sum_{m=1}^{J} exp((V_{it}^m(k))}$$

We also add an idiosyncratic stochastic term to the wage function and assume this to be normally distributed with mean 0 and variance σ_ω^2. The wage contribution to the likelihood function, conditional on unobserved heterogeneity (α^ω) for individual i in period t is given by:

$$l_{it}^{W} = \frac{1}{\sigma_\omega}\phi\left(\frac{W_{it} - \omega_{it}}{\sigma_\omega}\right)$$

where W_{it} is individual i's log earnings in period t and $\phi(.)$ is the standard normal probability density function.

5.1 The likelihood function

Combining the choice probabilities with the wage density, the overall likelihood contribution, conditional on unobserved heterogeneity given by α, is:

$$l_{it} = Pr(d_{it}^{j} = 1) \ * \left(l_{it}^{W}\right)^{I(d_{it}^{e}=1)}$$

The unconditional likelihood function is then given by:

$$logL(\Psi) = \sum_{i=1}^{N} ln\left(\sum_{k=1}^{K} \pi_k \prod_{t=1}^{T} l_{it}\right)$$

where π_k (the sample proportion of type k) is given by

$$\pi_k = \frac{exp(\theta^k)}{\sum_{n=1}^{K} exp(\theta^n)}$$

and where Ψ is a vector containing all parameters and θ^K is normalized to zero.[11]

5.2 Identification

As we observe all choices made, semester-by-semester, from the time of graduation at teacher-college, we avoid problems with initial conditions.[12]

[11] The model is estimated using R.

[12] There may be heterogeneity in the labor force experience among the teachers we sample prior to their enrollment in teacher college. We assume this has no impact on the subsequent choices made after

To estimate the parameters of the model, we rely on common distributional assumptions of the stochastic shocks. Standard instrumental variables techniques, which are generally based on outcomes and choices observed at a single point in time, do not apply to the framework here as it is not possible to obtain valid instruments for the sequences of outcomes and choices over time. We are interested in the effect of past choices on period-specific choices and wages, and this is achieved through estimation of a sequential choice model, which allows for a rich measure of multidimensional work experience that develops endogenously in the model.

The panel dimension of our data and the serial correlation in wages and choices help identify the time-invariant unobserved heterogeneity terms (α). Estimates may be inconsistent if the distributions of unobservables are misspecified. Choices and outcomes today depend on the past sequence of choices through the one-period time lags and the implied dynamic exclusion restrictions (Hu & Shum, 2012).

Our model also incorporates an exclusion restriction, i.e., a variable that is exogenous to the individuals but affects choices over time. In particular, local unemployment varies with time, place, and age, and introduces exogenous variation in outside options over time and place. Finally, the structure of the model delivers further exclusion restrictions, for instance by including separate work experience terms.

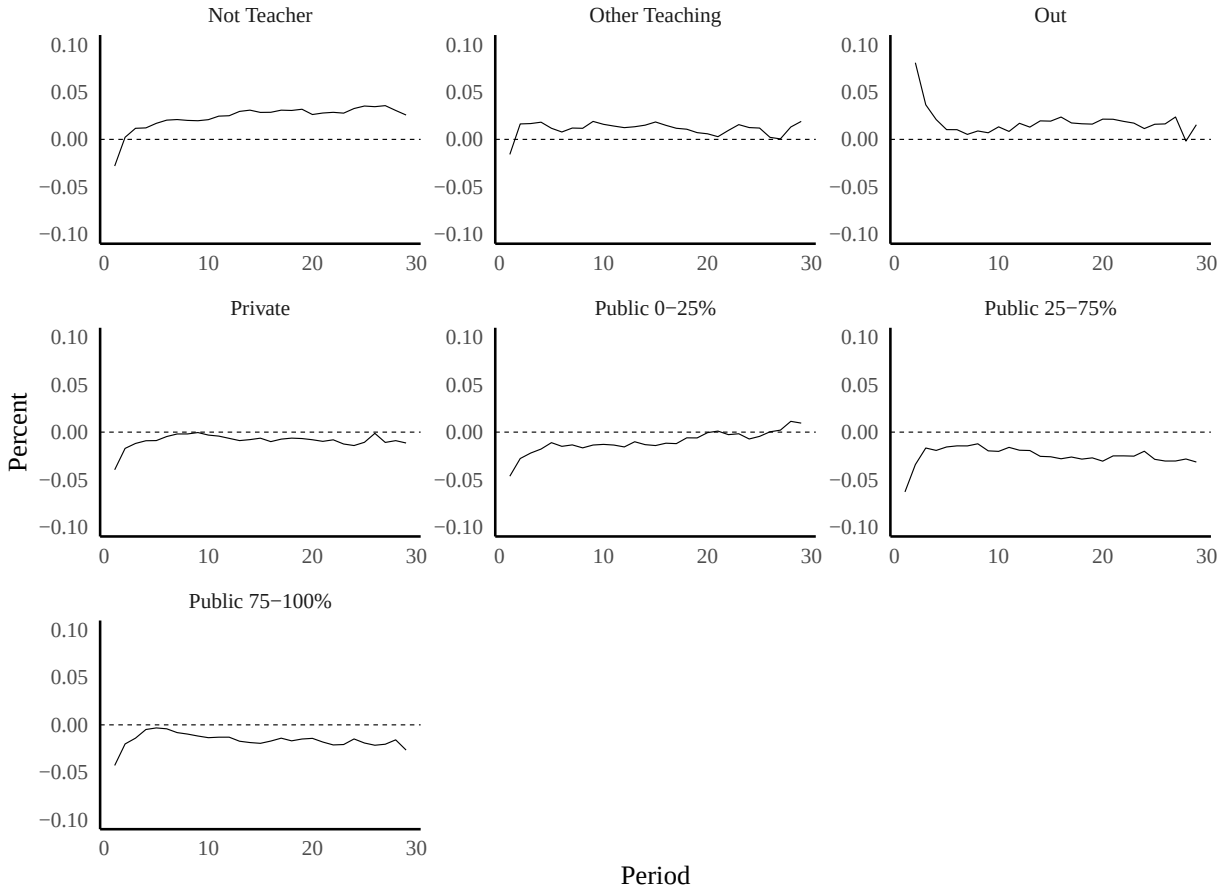
5.3 Validation

It is important to validate the models' ability to fit the data and the job choices made. The model formulation is intuitively appealing, but could nevertheless have been formulated differently, and as for any model the formulation chosen here partly reflects availability of data. For this reason, we validate the model by comparing actual choices with modeled choices.

college completion.

Figure 4 shows the percentage difference between predicted and actual choices made. The overall image is the the model fit appears to be very good. Nevertheless, a few caveats should be noted. In period 1, six of the seven options are about 5% too low. In contrast, the share of teachers in the "Out" option is too high compared to observed shares. This reflects that both the lagged choice and the accumulated experience terms, as they start at zero, cannot guide the models choice. After periods 1, this immediately improves. The option "Not Teacher" is 2-3 percentage-points too high in predicted share compared to actual. This mostly turns up in high and middle SES public schools having predicted shares that are lower than actual shares. The differences remain modest.

Figure 4: Model fit



Note: The figure shows the percentage difference in actual and predicted shares in each occupation across periods. Negative values indicate predicted share was lower than actually observed, and vice versa. The Y-axis is truncated at 0.10 in the "Out" panels first period.

6 Estimation results

6.1 Occupational choices

Ideally, the parameter estimates should be considered jointly, through simulations using all parameters (as we do below). Nevertheless, it is of interest to discuss some of the main features of the estimation results based on single point estimates alone.

Given the very large literature on teacher sorting and given the descriptive results above it comes as no surprise that we find that GPA from teacher college matters greatly for occupational choices, see Table 5. This demonstrates teacher sorting whereby high-quality teachers tend to teach in high SES public schools or in private schools.

However, given the model parametrization and the panel structure of our data, we can go some steps further. We introduce a parameter for the importance of GPA in the first 4 periods (2 years), a second parameter for the importance of GPA in periods 5-11, and a third parameter for experience after 6 years of experience (the top three rows of Table 5). Private schools can attract high-qualified teachers in periods 1-4 and 5-11, just out of college or with some experience, while GPA matters much less for private school occupation in the longer run, i.e. after period 11.

For the most attractive public schools, the top 25% in the socioeconomic distribution, a high GPA matters significantly in the first four periods, but not in subsequent periods. Schools with students of low SES background, the "public 0-25%"-option, attract teachers with low GPA, and over time this tendency does not wear off. Teachers with low GPA from teachers college have a prevalence of staying in low SES schools – partly reflecting their diminished outside options vis-à-vis teachers with higher GPA. Finally, note how high-quality teachers turn to the "other teaching"-option after some years (after period 11).

The parameter estimates for the first-order markov-terms, i.e., the one-period lags of occupational choice, reveal how there is substantial period-by-period dynamics across all types of occupation. The relative difference in the transitions within and between occupational options from one period to the next shows that there is a strong permanent component whereby teachers stay in their current type of occupation. Relative to the reference activity of being "outside", all other options are more prevalent. The highest parameter for staying in the same occupation group from period (t-1) to period t is found for "public 0-25%". Teachers from this type of occupation quite often transit to the middle SES public school option while, to a much lower degree, they transit "all the way" to a high SES public school.

A high level of local unemployment signals that outside options are less available. This makes the teachers take up an occupation in the low SES public school to a significantly higher degree than when unemployment is low and outside options therefore more available. This is also the case for the private school option. A similar finding is not seen for employment in one of the higher SES public school options. For the "public 75-100%"-option the point estimate even turns negative.

The model also includes an indicator-variable that takes the value 1 if the teacher is female and 0 if the teacher is male. The share of females completing teacher college is about 65% In addition to this high share of female teachers the parameter estimate for female teacher reveals how male teachers in particular are prone to choose occupations outside the public school. Male teachers are more often to be found in the "popular" teacher jobs, i.e. in the private school option or the high SES public school, and more often also outside elementary school. This corroborates findings by Podgursky et al. (2004) who show how female teachers are far more likely than males to teach. Being a male appears to enter positively in the job opportunity set as the most popular schools (private and high SES public) employ more males conditional on other characteristics

Table 5: Parameter Estimates for Occupational Choices

	Not Teacher	Other Teaching	Private	Public 0-25%	Public 25-75%	Public 75-100%
Teacher college GPA, periods 1-4	0.008	-0.019	0.149	-0.106	-0.026	0.145
	(0.037)	(0.03)	(0.035)***	(0.032)**	(0.029)	(0.032)***
Teacher college GPA, periods 5-11	0.027	0.009	0.121	-0.104	-0.023	0.011
	(0.03)	(0.025)	(0.031)***	(0.029)***	(0.025)	(0.028)
Teacher college GPA, periods 12-29	0.037	0.098	0.055	-0.101	-0.026	-0.067
	(0.034)	(0.029)***	(0.041)	(0.038)**	(0.032)	(0.037)
Female	-0.294	-0.24	-0.259	0.006	-0.016	-0.109
	(0.046)***	(0.038)***	(0.048)***	(0.046)	(0.04)	(0.045)*
Post reform	0.115	0.143	0.053	0.08	0.13	0.099
	(0.053)*	(0.043)***	(0.052)	(0.048)	(0.042)**	(0.047)*
Works part time	0.258	-0.204	-0.634	-0.73	-0.712	-0.617
	(0.044)***	(0.038)***	(0.049)***	(0.047)***	(0.04)***	(0.046)***
Unemployment rate, municipal level	-0.048	-0.009	0.02	0.031	0.014	-0.004
	(0.01)***	(0.008)	(0.01)*	(0.009)***	(0.008)	(0.009)
Not teacher	3.963	1.37	0.916	0.875	0.659	0.924
	(0.061)***	(0.075)***	(0.117)***	(0.121)***	(0.095)***	(0.109)***
Other teaching	1.616	3.875	1.165	1.486	1.238	1.399
	(0.071)***	(0.056)***	(0.093)***	(0.091)***	(0.072)***	(0.083)***
Private primary school	0.672	1.227	4.758	0.88	0.731	0.66
	(0.114)***	(0.084)***	(0.079)***	(0.118)***	(0.092)***	(0.109)***
Public primary school 0-25%	0.733	1.527	1.17	4.796	2.048	1.082
	(0.122)***	(0.084)***	(0.112)***	(0.086)***	(0.081)***	(0.105)***
Public primary school 25%-75%	0.778	1.408	1.06	2.186	4.123	1.64
	(0.093)***	(0.067)***	(0.093)***	(0.084)***	(0.061)***	(0.078)***
Public primary school 75-100%	0.557	1.269	0.919	0.992	1.466	4.425
	(0.106)***	(0.076)***	(0.103)***	(0.111)***	(0.076)***	(0.073)***
Sum of experience, Other teaching	0.01	0.146	-0.084	-0.074	-0.076	-0.111
	(0.008)	(0.006)***	(0.011)***	(0.01)***	(0.008)***	(0.01)***
Sum of experience, Out	-0.2	-0.342	-0.406	-0.485	-0.504	-0.504
	(0.008)***	(0.009)***	(0.015)***	(0.015)***	(0.012)***	(0.014)***
Sum of experience, Not teacher	0.215	-0.046	-0.114	-0.13	-0.12	-0.189
	(0.009)***	(0.011)***	(0.02)***	(0.02)***	(0.016)***	(0.021)***
Sum of experience, public schools 75%-100%	-0.01	0.000	-0.015	-0.091	-0.023	0.12
	(0.009)	(0.007)	(0.009)	(0.011)***	(0.007)***	(0.006)***
Sum of experience, public schools 25%-75%	-0.037	0.019	-0.013	-0.012	0.124	-0.001
	(0.009)***	(0.006)**	(0.008)	(0.007)	(0.006)***	(0.007)
Sum of experience, public schools 0%-25%	-0.011	0.014	-0.005	0.153	-0.007	-0.02
	(0.011)	(0.008)	(0.011)	(0.008)***	(0.008)	(0.01)
Sum of experience, Private school	-0.021	-0.016	0.127	-0.033	-0.036	-0.047
	(0.01)*	(0.008)*	(0.007)***	(0.01)***	(0.008)***	(0.01)***
intercept type 1	-1.684	-0.794	-0.402	-0.927	-0.429	-0.857
	(0.148)***	(0.129)***	(0.179)*	(0.164)***	(0.139)**	(0.149)***
intercept type 2	-1.065	0.003	-1.001	-1.359	-0.508	-0.629
	(0.163)***	(0.192)	(0.18)***	(0.177)***	(0.147)***	(0.162)***
intercept type 3	-0.729	-0.871	-1.549	-2.19	-1.188	-1.268
	(0.143)***	(0.139)***	(0.161)***	(0.166)***	(0.142)***	(0.158)***
intercept type 4	-1.327	-0.439	-0.899	-1.163	0.389	0.542
	(0.168)***	(0.135)**	(0.171)***	(0.161)***	(0.145)**	(0.177)**

such as GPA. Possibly, albeit somewhat speculative, because the schools seek to counter "feminization" of the school.

In Denmark there was a major reform of the public school coming to effect as of the school year 2014/2015 – following strikes in the spring of 2013. The "post-reform"-variable takes on a value of 1 in the time after this reform. The reform of the public school meant that public school teachers had much more fixed working hours and stricter working conditions with an increase in the teach-to-preparation time ratio. The reform is estimated to increase the share occupied outside elementary school but also increases the share working in high SES schools. The impact of the reform may appear relatively modest.

6.2 The importance of unobservable heterogeneity

Unobservables play an important role, and across the four discrete types there are major differences, which implies that some "types" have very strong preferences for certain occupational options over others. This is best illustrated by simulating the choices based on the estimated model parameters. Note from Table 6 how the estimated shares are relatively evenly spread over the four types, with shares ranging from 18 to 30 percent.

Table 6: Shares in unobservable type 1-4

Shares	Type
0.30	Share type 1
0.27	Share type 2
0.25	Share type 3
0.18	Share type 4

Note: Each individual type is computed based on Bayes formular and random draws from a uniform distribution.

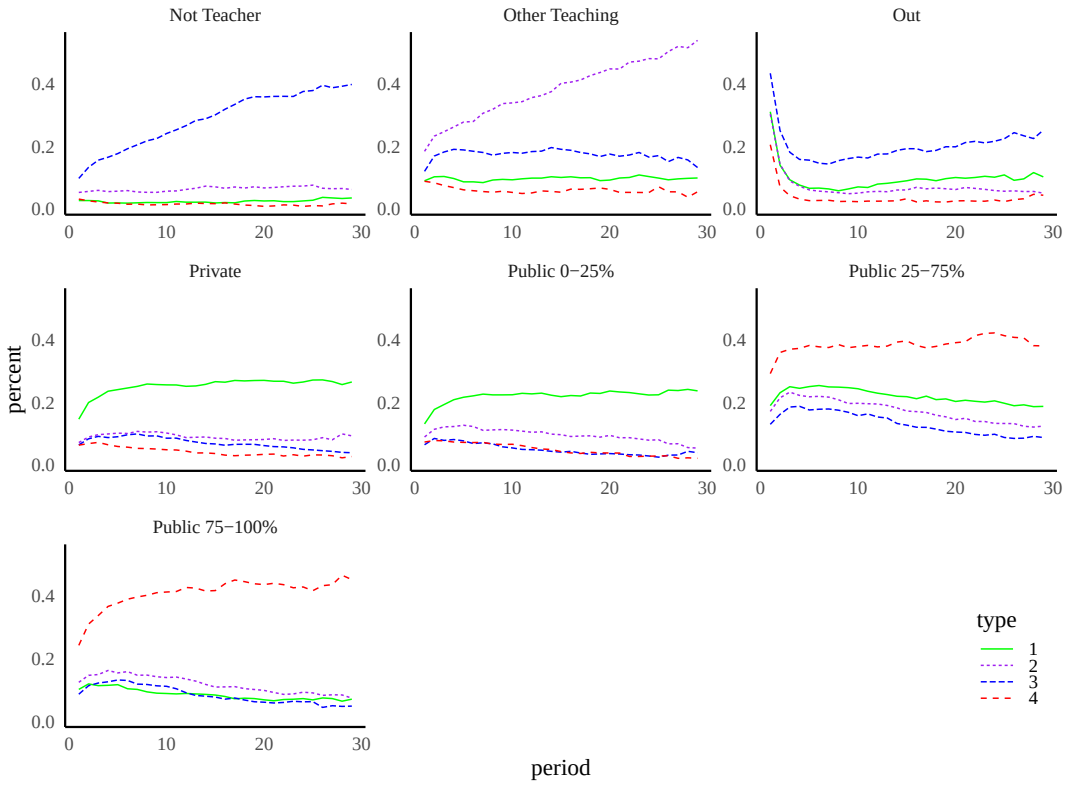
In Figure 5 we have simulated the choices made by teachers across periods and split by

the four unobservable "types". Evident from the figure, the unobservable types are very important for the choices made, and the types appear rather distinct from one another.

Type 1 (the green line) has the strongest preference for elementary school teaching via-à-vis the other three types. Type 2 (the purple line) favors "Other teaching"; type 3 (the blue line) favors not to be a teacher; and, type 4 (the red line) strongly favors the middle and high SES public schools.

These type-specific time-constant unobservable types are very important. How should we interpret them? They measure a variety of unobservable factors including personal individual traits such as motivation, drive, and perseverance. In principle, they can also be measuring many other factors that are not included as observable factors, say, whether the teacher has children. But it is difficult to think of how this type of variation would bring about these differences in preferences we see from our estimated type-parameters.

Figure 5: Occupational choices by type



7 Simulations

The model parameters are best understood by simulating choices. In addition, an advantage of the semi-structural approach is that it lends itself well to counterfactual simulations. Here, we simulate choices and wages across unobservable types and perform a counterfactual policy simulation.

7.1 Teacher quality and personality

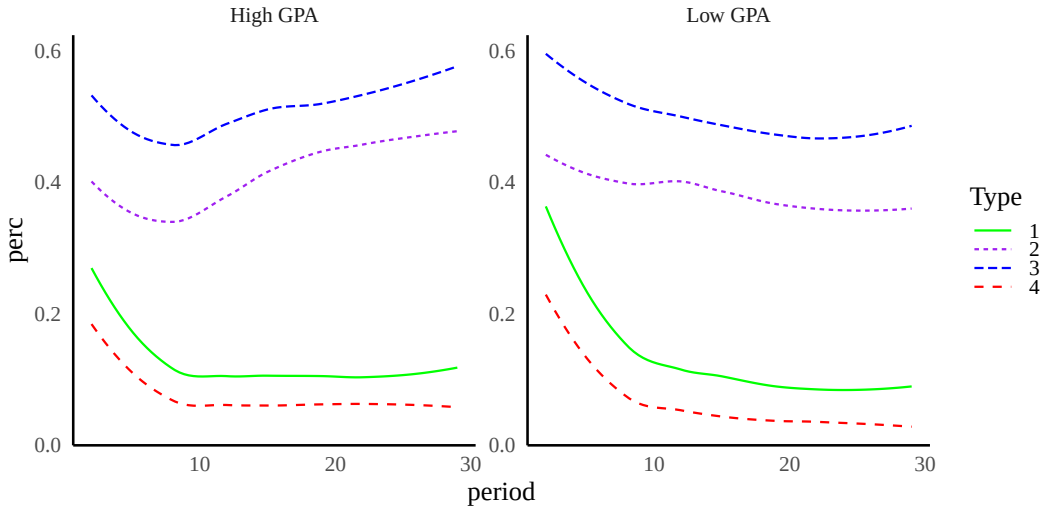
The existing literature has found that the best teachers (measured in a variety of ways) are more likely to leave low SES schools and even the teaching profession altogether (e.g. Podgursky et al., 2004; Rivkin et al., 2005). It is of great interest to analyze whether this also applies in our setting and the dynamics over time of any such development. In addition, given the unobservable types, it is also of interest to understand how such unobservables -in part measuring personality and unobserved preferences- alter the choices made compared to the importance of our teacher quality measure.

To this end, and in order to simplify the exposition, we combine "other teaching", "non-teaching", and "out" and name these choices "non-elementary". Given the shortage of qualified teachers this joint group is of special interest to analyze. In addition, following the above-mentioned studies, we also analyze how teachers choose low SES schools (0-25%) across the 4 different unobservable types, and by high and low college GPA.[13]

There are two main take-aways from figures 6 and 7. First, the difference in GPA is an important determinant for whether teachers are occupied in a low SES public school. Second, the unobservable types are otherwise oftentimes *more* important than GPA. The occupational patterns over time are highly determined by types even across very different

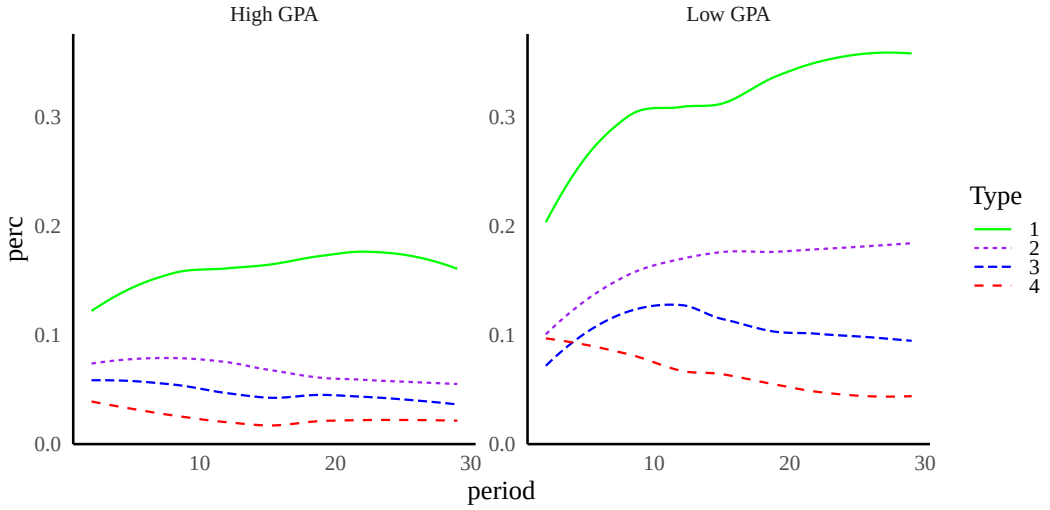
[13]High GPA is defined as 2 standard deviations above the mean and low GPA as two standard deviation below the mean. In other words, these are very large differences in GPA whereby we simulate a teacher from the very top with a teacher from the very bottom of the GPA distribution.

Figure 6: Share choosing non-elementary school, by high and low GPA and by type



Note: High (low) GPA is defined as 2 standard deviations above (below) the mean.

Figure 7: Share choosing low SES public school, by high and low GPA and by type



Note: High (low) GPA is defined as 2 standard deviations above (below) the mean.

levels of GPA. Especially the choice of working outside elementary school appears to be driven by unobservable time-constant (personal) differences rather than GPA. Even though we simulate the model for very large differences in GPA.[14]

[14]In the appendix, Figure 10 confirm that this is also very much the case for mid and high SES public

7.2 Wage returns to leaving teaching

In our estimation routine we jointly maximize the likelihood of all parameters including a wage equation in which we measure the returns to the occupational choices made. In order to illustrate the parameters we simulate, by level of college GPA, the wage profile of a teacher who start out with 9 periods in the middle SES public sector.[15] In our wage simulation we compare the wage profile between two distinct choices made after 9 periods in a middle SES public school. Either the teacher stays in the middle SES public school until period 29 or the teacher chooses the "not teacher" option.

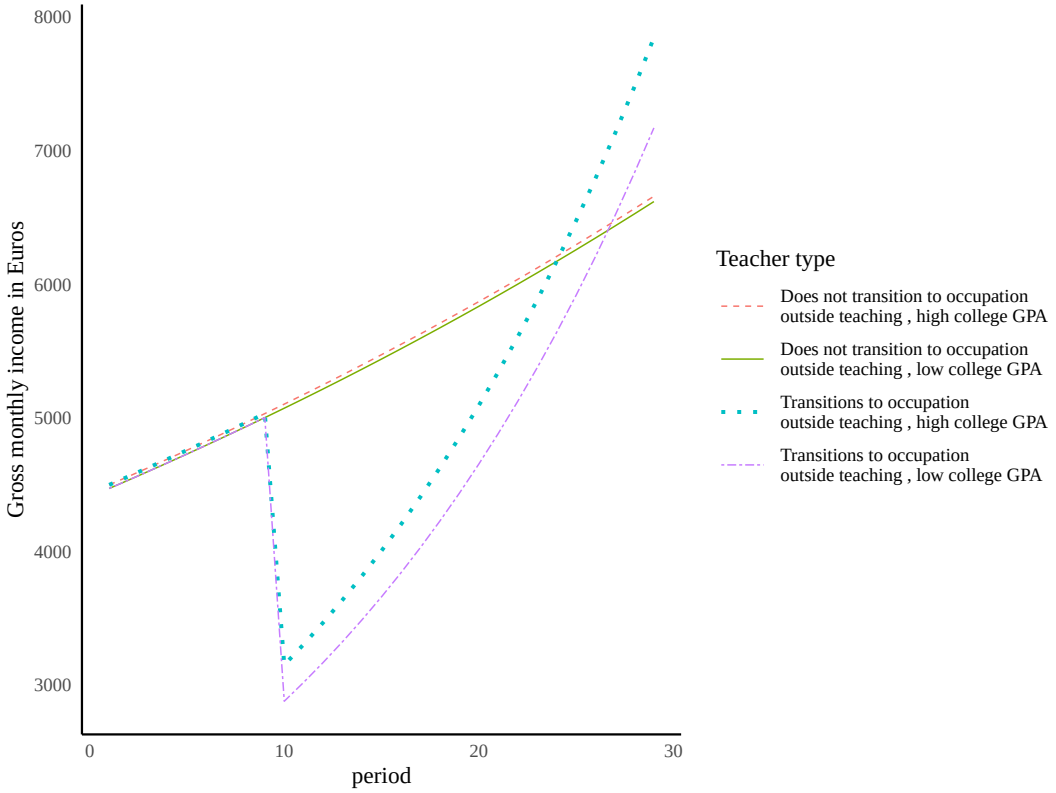
The simulation is split by a high quality teacher, i.e. a teacher with a college GPA which is 1 standard deviation above the average, and a low quality teacher with GPA 1 standard deviation below the average. According to our wage parameter estimates, the quality difference has no impact for teachers in the public school. For teachers choosing the "Not teaching" work option, the wage is slightly higher for a high GPA individual than for a low GPA individual. But the main point from Figure 8 is that there is an immediate and very large wage penalty for choosing a job outside teaching. After some 15 periods this is reversed. It is staggering how the 10% of our sample, who choose the "Not teaching" work option, substantially reduce their wage earnings. The utility loss of being a teacher has to out-weigh the loss in earnings.

Scafidi et al. (2006) examine the extent to which teachers leave teaching for higher paying jobs. They find that very few of those who leave teaching take jobs that pay more than their salary as a teacher. Our results corroborate this finding.

schools and private schools, although some differences related to GPA-levels also do appear.

[15] A table with wage parameter estimates is included in the appendix.

Figure 8: Wage path simulations



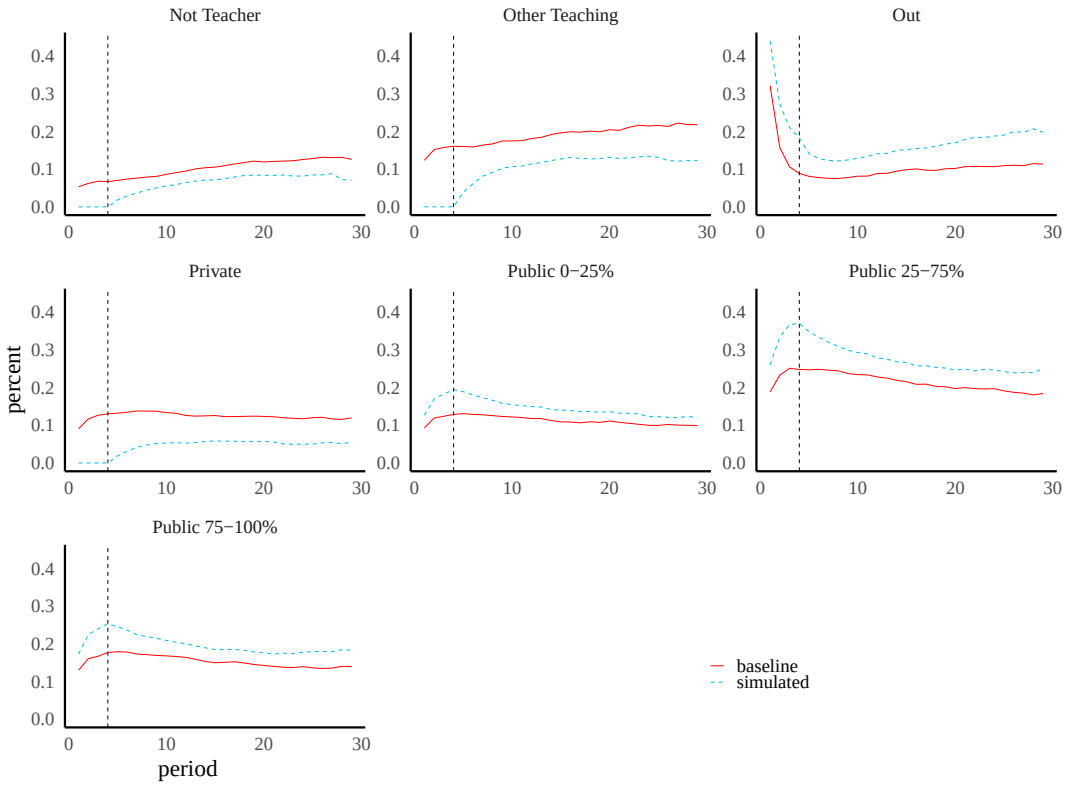
Note: In periods 1-9 the choices are middle SES public school. In period 10 and until period 29 we see the wage path if the teacher choose the "Not teacher"-option. The two colors signify a 1 standard deviation higher (lower) than average GPA.

7.3 A grace period in the public school system

Here, we simulate the impact of a hypothetical policy whereby newly graduated teachers are to abide to a grace period of 2 years in which they have to teach in the public school system. The only other alternative we allow in the simulation is the "other" (outside) option, which is the reference group. This policy will induce an increase in the accumulation of public school experience during the first 4 periods. We are then mostly interested in how such a grace period wears off as time passes and how it affects teacher sorting and hence the quality of teachers, say, in the 0-25% SES public school.[16]

[16]This hypothetical policy may appear unlikely to be introduced in the real world. Note though, that it resembles the policy found for those who enter military pilot training in Denmark. They sign a contract

Figure 9: Counterfactual simulation of a public school grace period



Note: The baseline shows the shares in each occupation in total (across the four types). The dashed light-blue line shows the choices when a grace period of 4 periods is introduced.

In Figure 9 we show the results of this thought-experiment. By definition, the share in "not teacher", "other teaching", and "private school" are set to zero in the first 4 periods. This has a permanent impact on subsequent choices, and in this manner it does result in a permanent increase in teachers in the public school system. However, the improvement is lowest in the low SES public school – both during the 4 periods but also after. In addition, the reference group "out" also increases and it is noteworthy that the initial increase in the "out"-option even increases as time passes. Based on this simulation, a grace period has mixed effects and very little impact on low SES public schools.

with the air force so that, after completing up to 8 years of education, they have to serve 12 years in the air force after completing the education.

8 Discussion

This paper is motivated by teacher sorting and the lack of qualified teachers – especially for students in low SES public schools. While the analysis has revealed several new findings concerning the dynamics (or lack of dynamics) over the first 15 years of teachers' careers, the results generally find support in previous research.

Raw data inform us that 40% of the college educated teachers work outside primary and lower secondary school 15 years after graduation. Our results show that teachers who withdraw from the elementary school especially are to be found among the unobserved discrete "types" 2 and 3, who constitute about half of the sample. These are teachers who already in period 1 have a relatively high share outside the public or private schools, and this share increases steadily as time passes and experience increases. For the other half of the sample, types 1 and 4, the outside options are much less preferred. Both initially and over time.

Given that such time-constant, possibly inherent, preferences determine teachers choices, an effective policy-response, i.e. an initiative that would make these teacher types enter/stay in the public or private school system, is a challenge. The occupational patterns do not result in improved wage outcomes for those who choose not to teach in the public/private schools. In other words, differential wage-opportunities do not appear to have a major impact. However, even though half of the teachers, types 2 and 3, may opt out of the elementary school most of them do enter the teaching profession. On one hand, our results point towards initiatives supporting novice teachers better in order to improve retention. Mentoring would be one such initiative. This follows previous research, such as Bruno et al. (2020) who find that for novice teachers' professional contexts may have important implications for their development and retention. On the other hand, teacher types 3 and 4 hardly enter low SES public schools, so their share in low SES public schools does not diminish much with time after graduation. For these teachers, who constitute

more than 40%, supporting initiatives would make little difference, at least in the short term.

The results conveyed in Nguyen et al. (2020) also points towards mentoring and introductory support has the potential to significantly improve teacher retention. Our results serve to explain why such support-initiatives appear effective: they are targeted novice teachers and the early career, the first 2-5 years, are highly determinant of later occupational choices.

8.1 Concluding remarks

Teacher quality differ within the public school and between public and private schools. The best-qualified teachers, as measured by their GPA from teacher college, sort towards high SES public schools and towards private schools. This sorting occurs also for novice teachers with no or little experience. Yet, either immediately or after only a few years, many of the college educated teachers do not work as a teacher in the public or private school system. After 15 years, 40% of the teachers do not teach in primary or lower secondary school. While the literature to date mostly concludes that the most academically prepared teachers–measured by, e.g., ACT scores or college degree–are more likely to leave both low SES public schools or the teaching profession our results only partially corroborate with these results.

Unobservables – in part measuring personality traits – are generally at least as important as GPA in explaining who leaves low SES public schools or elementary school.

To the best of our knowledge, this study is the first to demonstrate teacher sorting both between public, private and outside occupational options while also differentiating within the public school system.

Given our parameter estimates, time-constant unobservable "types" is a key explanation. The differences in teacher quality are not very pronounced within the public schools

with the important exception that low GPA teachers are much more likely to work in low-SES public schools.

The policy implications include a special focus on novice teachers. Teacher careers are often settled relatively soon after graduation and with a high level of inertia after, say, 6-8 years.

Across academic disciplines, there is an overwhelming amount of research on the relationship between students, teachers and schools, and the importance of teacher and school quality. Nevertheless, there only very few studies use a structural dynamic approach. We believe such models to be useful and encourage further developments. Including data on personality traits appears particularly relevant.

9 References

AE Rådet (2023). Næsten hver femte lærer i folkeskolen har ikke en læreruddannelse (In Danish). Analyse, AE Rådet, Copenhagen.

Ashworth, J., V. J. Hotz, A. Maurel & T. Ransom (2021). Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences. *Journal of Labor Economics*, **39**(4): 931-964.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., Wyckoff, J. (2008). Who Leaves? Teacher Attrition and Student Achievement, NBER 14022, National Bureau of Economic Research.

Boyd, D., Lankford, H., Loeb, S., Wyckoff, J. (2013). Analyzing the determinants of the matching of public school teachers to jobs: disentangling the preferences of teachers and employers. *Journal of Labor Economic*s, **31**(1): 83-117.

Behrman, J.R., M.M. Tincani, P.E. Todd, K.I. Wolpin (2016). Teacher Quality in Public and Private Schools under a Voucher System: The Case of Chile. *Journal of Labor Economics*, 2016, **34**(2), pt.1: 319-362.

Bernal, R. & M. P. Keane (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*,**156**: 164-189.

Cameron, S. V. & J.J. Heckman (1998). Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy,* **106**(2): 262-333.

Clotfelter, C. H. F. Ladd, J. Vigdor & J. Wheeler (2007). High-Poverty Schools and the Distribution of Teachers and Principals, 85 N.C. L. Rev. 1345.

Eckstein, Z. & K. E. Wolpin (1999). Why Youth Drop Out of High Scool: The Impact of Preferences, Opportunities, and Abilities. *Econometrica*, **67**(6): 1295-1339.

Eriksson, T. & N. Kristensen (2014). Wages or Fringes? Some Evidence on Trade-Offs and Sorting. *Journal of Labor Economics*, **32**(4): 899-928.

EU (2010): Early School Leaving. Lessons from research for policy makers. European Commission, 2010.

Gensowski, M., R. Landersø, D. Bleses, P. Dale, A. Hœjen & L. Justice (2024). Public and Parental Investments, and Childrens Skill Formation. IZA Discussion Paper No. 16956.

Geweke, J. & M. Keane (2000). Bayesian inference for dynamic discrete choice models without the need for dynamic programming. In *Simulation-Based Inference in Econometrics: Methods and Applications,* R. S. Mariano, T. Schuermann, and M. Weeks, (eds.): 100–131, Cambridge University Press, Cambridge.

Goldhaber, D., Lavery, L., Theobald, R., 2015. Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*,**44** (5): 293-307.

Goldhaber, D., M. Kasman, V. Quince, R. Theobald & M. Wolff (2023). How did it get this way? Disentangling the sources of teacher quality gaps through agent-based modeling. *Social Science Research*, **116**, 102941: 1-12.

Hansen, J. & N. Kristensen (2024). The bottom 20 percent: early career paths of youth with low grade point average. *Scandinavian Journal of Economics*, In print.

Hanushek, E., J. Kain & S. Rivkin (2004). Why public schools lose teachers, *Journal of Human Resources*, **39**(2): 623-54.

Hanushek, E. & S. Rivkin (2010).The Quality and Distribution of Teachers under the No Child Left Behind Act. *Journal of Economic Perspectives*, **24** (3): 133-150.
Heckman, J.J. & B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, **52**: 271-320.

Holmlund, H., M. Lindahl & E. Plug (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, **49**(3): 615-651.

Hu, Y. & M. Shum (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics*, **171**: 32-44.

Isenberg, E., J. Max, P. Gleason, J. Deutsch (2022). Do low-income students have equal access to effective teachers? Evidence from 26 districts. *Educational Evaluation and Policy Analysis* **44**(2): 234-256.

James, J. & J. Wyckoff (2022). School Segregation, Teacher Sorting, and the Distribution of Teachers. EdWorkingPaper: 22-659. https://doi.org/10.26300/tv5x-6t2.

Jørgensen, T. & M. Tô (2018). Robust Estimation of Finitite Horizon Economic Models, *Computational Economics*, forthcoming.

Keane, M. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics*, **156**: 3–20.

Keane, M. & K. L. Wolpin (1997). The career decisions of young men. *Journal of Political Economy*, **105**(3): 473–522.

Kristensen, N. (2023). Om manglen på uddannede folkeskolelærere. Notat (In Danish). VIVE.

Kristensen, N. (2024). The Importance of Educated Public School Teachers for Children at Risk of Dyslexia. Manuscript, VIVE.

Lazear, E. P. (1998). Personnel economics for managers. New York: Wiley.

Lazear, E. P. (2005). Output-based pay: Incentives or sorting? *Research in Labor Economics*, **23**: 1-2.

Lazear, E. P. & K. Shaw (2007). Personnel economics: The economistâs view of human capital. *Journal of Economic Perspectives*, **21**(4): 91-114.

Lazear, E. P. & P. Oyer (2013). Personnel economics. In *Handbook of organizational economics*, ed. Robert Gibbons and John Roberts, 479-519. Princeton, NJ: Princeton University Press.

Lindqvist, P. & U. K. Nordanger (2016). Already elsewhere A study of (skilled) teachers' choice to leave teaching. *Teaching and Teacher Education*, **54**:88-97.

Marra, M.(2021). Confessions of a Public School Teacher: A 35-Year Veteran's Assessment on How Public Education Can Make the Grade. *Dorrance Publishing Co.*

Mortensen, D. T. & C. A. Pissarides (1999). New developments in models of search in the labor market. Chapter 39 in Orley C. Ashenfelter and David Card (eds.) Handbook of Labor Economics, **3**, Part B: 2567-2627.

Nguyen, T.D., L.D. Pham, M. Crouch, & M.G. Springer (2020). The correlates of teacher turnover: An updated and expanded Meta-analysis of the literature. *Educational Research Review*, **31**, 100355: 1-17.

OECD (2024), Education at a Glance 2024: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/c00cad36-en.

Oyer, P. & S. Schaefer (2005). Why do some firms give stock options to all employees? An empirical examination of alternative theories. *Journal of Financial Economics*, **76**(1): 99-133.

Podgursky, M., R. Monroe & D. Watson (2004). The academic quality of public school

teachers: an analysis of entry and exit behavior. *Economics of Education Review*, **23**:507-518.

Rodriguez, L.A., T.D. Nguyen & M.G. Springer (2023). Revisiting Teaching Quality Gaps: Urbanicity and Disparities in Access to High-Quality Teachers across Tennessee. *Urban Education*: 1-38.

Rust, J. (1994). Structural Estimation of Markov Decision Processes. In *Handbook of Econometrics*, Volume 4, edited by Robert F. Engle and Daniel L. McFadden, 3081–3143. Amsterdam and San Diego: Elsevier, North-Holland.

Rust, J. (2014). The Limits of Inference *with* Theory: A Review of Wolpin (2013). *Journal of Economic Literature*, **52**(3): 820–850.

Salop, J. & S. Salop (1976). Self-selection and turnover in the labor market. *Quarterly Journal of Economics*, **90**(4): 619-27.

Scafidi, B., D.L. Sjoquist & T.R. Stinebrickner (2006). Do Teachers Really Leave for Higher Paying Jobs in Alternative Occupations? *B.E. Journal of Economic Analysis & Policy*, Advances, **6**(1).

Willis, R. J. & S. Rosen (1979). Education and Self-Selection. *Journal of Political Economy*, **87**(5), pt.2: s7-s36.

Rivkin, S. G., E. A. Hanushek & J. F. Kain (2005). Teachers, Schools, and Academic Achievement. *Econometrica* **73**(2): 417-58.

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, **100**: 61-78.

Wolpin, K.I. (1992). The Determinants of Black-White Differences in Early Employment Careers: Search, Quits, Layoffs and Endogenous Wage Growth. *Journal of Political Economy*, **100**(3): 535-560.

Appendix: Supplementary tables and figures

Table 7: Schools Characteristics by 9th grade GPA deciles

	Deciles									
	1	2	3	4	5	6	7	8	9	10
Students avg. GPA in 9th grade	5.4	6.1	6.4	6.6	6.8	7.0	7.2	7.4	7.7	8.2
Teachers avg. GPA from teacher college	7.0	7.1	7.2	7.2	7.3	7.4	7.6	7.7	7.8	7.9
Parents avg. income (euro)	29,521	30,587	31,760	32,090	32,686	33,543	34,630	36,705	38,824	42,971
Parents avg. lenght of education	13.1	13.5	13.8	13.9	14.2	14.3	14.5	14.9	15.3	15.6

Table 8: Transition matrix across periods

Status in t-1	Transition matrix in periods 0 - 5							Transition matrix in periods 6 - 10						
	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out
Public 0-25%	0.82	0.08	0.02	0.02	0.03	0.01	0.02	0.87	0.07	0.01	0.01	0.03	0.00	0.01
Public 25-75%	0.04	0.84	0.04	0.01	0.03	0.01	0.03	0.04	0.88	0.03	0.01	0.02	0.00	0.02
Public 75-100%	0.01	0.05	0.86	0.01	0.03	0.01	0.03	0.01	0.04	0.90	0.01	0.02	0.00	0.02
Private	0.01	0.02	0.01	0.89	0.03	0.01	0.02	0.01	0.02	0.01	0.93	0.02	0.00	0.02
Other Teaching	0.04	0.06	0.04	0.02	0.77	0.03	0.04	0.02	0.04	0.02	0.01	0.85	0.03	0.03
Not Teacher	0.03	0.06	0.04	0.03	0.09	0.68	0.07	0.01	0.02	0.01	0.01	0.06	0.83	0.06
Out	0.05	0.12	0.07	0.06	0.08	0.05	0.57	0.03	0.06	0.04	0.03	0.08	0.06	0.69

Status in t-1	Transition matrix in periods 11 - 15							Transition matrix in periods 16 - 20						
	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out	Public 0-25%	Public 25-75%	Public 75-100%	Private	Other Teaching	Not Teacher	Out
Public 0-25%	0.89	0.05	0.01	0.01	0.03	0.00	0.01	0.89	0.06	0.01	0.01	0.02	0.00	0.01
Public 25-75%	0.02	0.90	0.02	0.01	0.02	0.00	0.01	0.02	0.91	0.02	0.01	0.02	0.00	0.01
Public 75-100%	0.00	0.03	0.91	0.01	0.02	0.00	0.02	0.00	0.03	0.91	0.01	0.02	0.01	0.01
Private	0.00	0.01	0.01	0.93	0.02	0.00	0.02	0.00	0.01	0.01	0.94	0.02	0.00	0.01
Other Teaching	0.01	0.02	0.01	0.01	0.89	0.03	0.03	0.01	0.02	0.01	0.01	0.91	0.02	0.02
Not Teacher	0.01	0.01	0.01	0.01	0.04	0.87	0.05	0.00	0.01	0.01	0.01	0.03	0.91	0.04
Out	0.02	0.03	0.02	0.02	0.07	0.06	0.77	0.01	0.02	0.02	0.02	0.06	0.05	0.82

Note: Transitions from t-1 to t in total.

Table 9: Parameter Estimates for Wage Equation

	Parameter estimate
Teacher college GPA, periods 1-4	0.003
	(0.004)
Female	-0.057
	(0.008)***
Post reform	0.025
	(0.01)**
Works part time	-0.233
	(0.008)***
Unemployment rate, municipal level	0.003
	(0.002)
Not teacher	25.804
	(0.02)***
Other teaching	26.189
	(0.016)***
Private primary school	26.303
	(0.018)***
Public primary school 0-25%	26.373
	(0.018)***
Public primary school 25%-75%	26.362
	(0.015)***
Public primary school 75-100%	26.377
	(0.017)***
Sum of experience, Other teaching	0.028
	(0.001)***
Sum of experience, Out	-0.013
	(0.002)***
Sum of experience, Not teacher	0.048
	(0.002)***
Sum of experience, public schools 75%-100%	0.014
	(0.001)***
Sum of experience, public schools 25%-75%	0.014
	(0.001)***
Sum of experience, public schools 0%-25%	0.014
	(0.001)***
Sum of experience, Private school	0.015
	(0.001)***
Teacher college GPAXNot Teacher	0.042
	(0.012)***
Intercept type 1	-15.974
	(0.018)***
Intercept type 2	-15.969
	(0.018)***
Intercept type 3	-15.982
	(0.018)***
Intercept type 4	-15.968
	(0.019)***

Figure 10: Quality and type: Private schools and medium and high SES public schools