The Rise of Women and Decline of Education

Ann-Sofie Kolm* Per Krusell† Kieran Larkin‡ Kurt Mitman§

21 March 2025

Abstract

This paper investigates how a decline in barriers to market production facing women has affected female labor force participation, occupation decisions, and aggregate human capital formation. Since the 1960s, female labor force participation has increased, the share of college-educated women in the labor force working as teachers has fallen, and average teacher quality has declined. We investigate the implications for the aggregate economy of reallocating highly productive women from teaching and into the market sector. Teaching is unusual because the government fixes the contract most teachers face and does not fully reflect differences in teacher productivity. We interpret these shifts in U.S. and Swedish data using a framework incorporating declining labor market barriers and inefficient teacher contracts. We derive conditions under which this can distort the first best allocation, even if the planner can choose the contract instruments optimally. Our quantitative analysis suggests that lower barriers to women's labor market entry explain approximately 60% of the observed decline in teacher quality, resulting in diminished aggregate human capital. Furthermore, we show that inefficiencies from suboptimal teacher contracts exacerbate misallocation, causing additional welfare losses of almost 4%. Despite these trends, reallocating talent away from education towards other sectors generates significant overall welfare gains.

Keywords: Female participation, human capital, occupational choice, teachers, misallocation, government wages

JEL Codes: J16, J24, I20, D70

*ann-sofie.kolm@su.se.
†per.krusell@iies.su.se.
‡kieran.larkin@iies.su.se.
§kurt.mitman@iies.su.se.

1 Introduction

The last 60 years have seen a remarkable increase in female labor force participation across the developed world (Goldin, 1990). In the US, women's labor force participation increased from 44.5 percent in the 1960s to 77.6 percent by the 2000s. In Sweden, women's participation rate rose from around 55 percent in the 1960s to over 80 percent by the 1990s. While the participation rate has been fairly stable over the last 30 years, the share of women in traditionally male-dominated occupations is still rising. This is particularly true for high-skilled occupations, suggesting the transition to the full integration of women into the labor market is still incomplete (Goldin, 2014).[1]

The substantial rise in women's labor force participation has coincided with a marked decline in educational outcomes, as illustrated by decreasing student test scores and declining indicators of teacher quality. Historically, female teachers had higher average educational attainment than women employed in other sectors. However, this educational gap has narrowed significantly, driven by a rapid increase in the educational levels of women in non-teaching occupations. Additionally, alternative indicators of teacher quality, such as aptitude tests and standardized scores among teaching candidates, have also declined substantially over recent decades, particularly in Sweden and other advanced economies (Grönqvist and Vlachos, 2016). Figure 1 illustrates this broader trend by depicting the decline in mathematics scores on the PISA assessment across selected OECD countries from 2003 to 2021, during which the average OECD score decreased from 502 to 480.

We argue that these developments are fundamentally interconnected. Specifically, the reallocation of high-skilled women into the market sector has had an overlooked consequence: a *negative* impact on the average quality of teachers. This reallocation of talent has significant implications for overall market outcomes due to teachers' central role in shaping human capital. Teacher quality, or value-added, has been shown to have an important causal impact on student outcomes later in life (See: Rivkin et al. (2005), Hanushek and Rivkin (2012), Chetty et al. (2014), Chetty et al. (2016)). To put it succinctly, the quality of teachers matters.

In this paper, we examine how reduced barriers to labor market participation for highly skilled women have affected talent allocation between market occupations and teaching, as well as the aggregate economic consequences of these shifts. We hypothesize that teaching previously attracted talented women due to limited access to other high-skilled professions.[2] Today, however, women increasingly select alternative high-skilled careers,

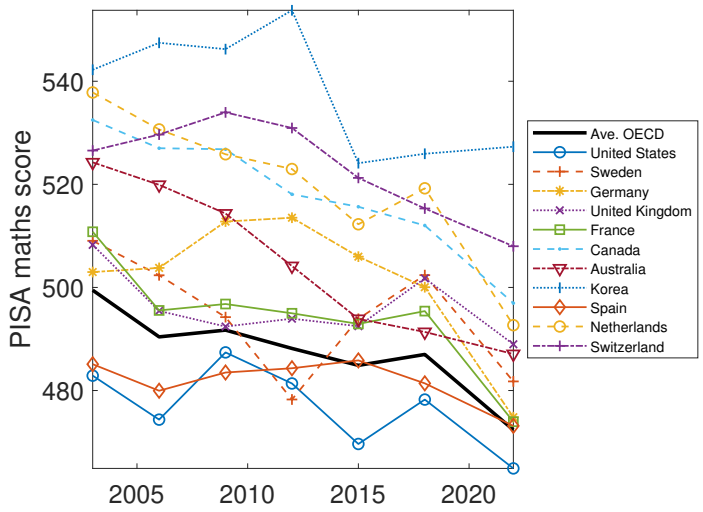
[1] As one example, in Sweden the share of women among those in higher education getting a degree in business, economics, political science, and law, has increased from 25 percent in 1975 to around 60 percent in 2020 (Halldén and Stenberg (2023)).

[2] Looking back over 60 years, while women's participation in most market sectors was limited, teaching was an exception. In the United States during the 1960s, around 45 percent of college-educated women participating in the labor force

such as medicine or law, instead of teaching. Consequently, in productivity terms, the mid-20th-century teacher workforce was arguably "too high" in quality relative to what would have been expected without labor market barriers for women. While the overall allocation of talent has improved as women enter various high-skilled market occupations, this shift may come with the trade-off of having teachers who, on average, possess lower levels of talent and productivity.

A further consideration is that the allocation of talent between teaching and market sectors may be distorted due to rigidities inherent in teacher contracts. In most countries, the teaching profession is subject to substantial regulation, with employment conditions typically set by the government and frequently shaped by union negotiations. This regulatory framework often results in relatively uniform pay structures for teachers, which rarely align fully with individual productivity differences. Consequently, these contractual arrangements may introduce inefficiencies, creating an additional distortion between the optimal allocation of talent and the equilibrium we observe in practice.

Figure 1: Decline in PISA scores



Notes:

In this paper, we document a set of empirical facts consistent with falling teacher quality. We show that in the US and Swedish microdata, the average quality of teachers has fallen over time. Further, the most highly productive female teachers have reallocated to work in market activities. We then take a structural approach to studying the teacher allocation problem. We build an occupational choice model featuring correlated productivities

were employed as teachers, a figure that dropped to approximately 18 percent by the 2000s. By comparison, only about 8 percent of college-educated employed men worked as teachers. This gendered occupational distribution can be attributed to historical regulations and social norms. In Sweden, starting as early as the mid-1800s, married women were restricted from paid employment unless they worked as teachers. As employment opportunities gradually expanded for women in the early 1900s, teaching remained a primary avenue for female employment (Molinder and Pihl, 2021). Consequently, the proportion of female teachers has historically been high; over the past four decades in Sweden, women comprised roughly 60 percent of qualified teachers in middle and high schools and over 80 percent in pre-school and elementary education.

between teaching and market productivity, human capital production, an occupation decision, an educational attainment choice, and a social planner that optimally chooses a teaching contract subject to constraints.

In a simplified version of the model, we first characterize the solution to the social planner's problem, taking as given the barriers that women face. Next, we compare that solution to a constrained planner's problem, where the contract that the planner can design to offer to teachers features only a wage level and minimum skill cut-off. The particular restriction on the instruments the planner can use is motivated by the empirical fact that teacher wages tend to be highly regulated, and entry to the teaching occupation typically involves taking tests to achieve certification. If productivity is perfectly correlated between teaching and market occupations, the constrained planning problem will still be able to achieve the first best allocation. When productivities are positively, but not perfectly, correlated, the allocation in the constrained planning problem features a distortion. Some women with high productivity in teaching are instead allocated to the market (compared to being allocated to teaching in the first-best), while some women with medium productivity for market occupations end up as teachers, even though they would be working in the market sector in the first best allocation.

Next, we calibrate the model to be consistent with salient facts on female labor force participation, occupational choice, teacher value-added, and wage dispersion. We use the model to quantify the effect of the fall in barriers to market production facing women between 1964 and 2020, using the observed allocations across teaching and non-teaching occupations. We assume in each period the social planner can optimally choose a teaching contract that maximizes welfare, subject to the constraint that it features only a fixed wage and a minimum skill requirement. The model provides predictions for 1.) the increase in aggregate output and welfare due to the improved allocation of talents across occupations as women entered the labor market; 2) the extent to which this re-allocation can explain the observed reduction in teacher quality and student performance, 3.) the cost imposed by the misallocation due to contracting frictions and how it has changed over time.

We find that the barriers to market production facing women have fallen substantially over time, declining by almost 60 percent between 1970 and 2000. However, relative to men a significant productivity penalty remains. This effective increase in female productivity has greatly raised welfare. In this period, the optimal contract features a declining minimum skill cutoff and U-shaped teacher wage. Fitting the data, the share of females working that are teachers has fallen and the average teacher skill level has declined. This has reduced human capital production in the teaching sector by around 18 percent. As a result, male output has fallen. This may explain declining outcomes for low-skilled men in the labor market as has been recently documented Autor and Dorn (2013). With a fixed wage, the model reproduces both the decline in the average teacher skill level and around half the fall in the teacher wage premium observed in the US data. Restrictions on the

teacher contract have a fairly small welfare cost, with the planner able to implement an allocation close to the first best. The welfare deviation from the optimal teacher contract is about 1 percent. Through the lens of the model, the reason for this is the strong substitutability across teacher types. However, it conforms with the observation that many education systems have increased the use of lower-skilled ancillary teachers (e.g. Helen et al. (2018)). One exception to the welfare result is in the case of extremely rigid contracts. If the planner cannot vary *either* instrument after optimally choosing the contract in the initial period, high-productivity teachers reallocating to the market sector are not replaced with lower-quality substitutes. In this case, there is a large fall in human capital and a substantial welfare loss occurs. Under this contract, welfare is 8 percent lower than it would be otherwise in 2020. Our work is related to a recent paper by Alder et al. (2023) who also focus on the consequences of occupational choice for human capital. We differ by considering the restricted nature of teacher contracts and explicitly solving for the constrained social planner's problem. Both papers are complementary.

The roadmap for the rest of the paper is as follows: in Section 2 we document motivating empirical facts; in Section 3 we set out the model; in Section 4 we consider a benchmark case with perfectly correlated abilities; in Section 5 we set out the model calibration; in Section 6 we undertake the main quantitative experiment; while Section 7 concludes.

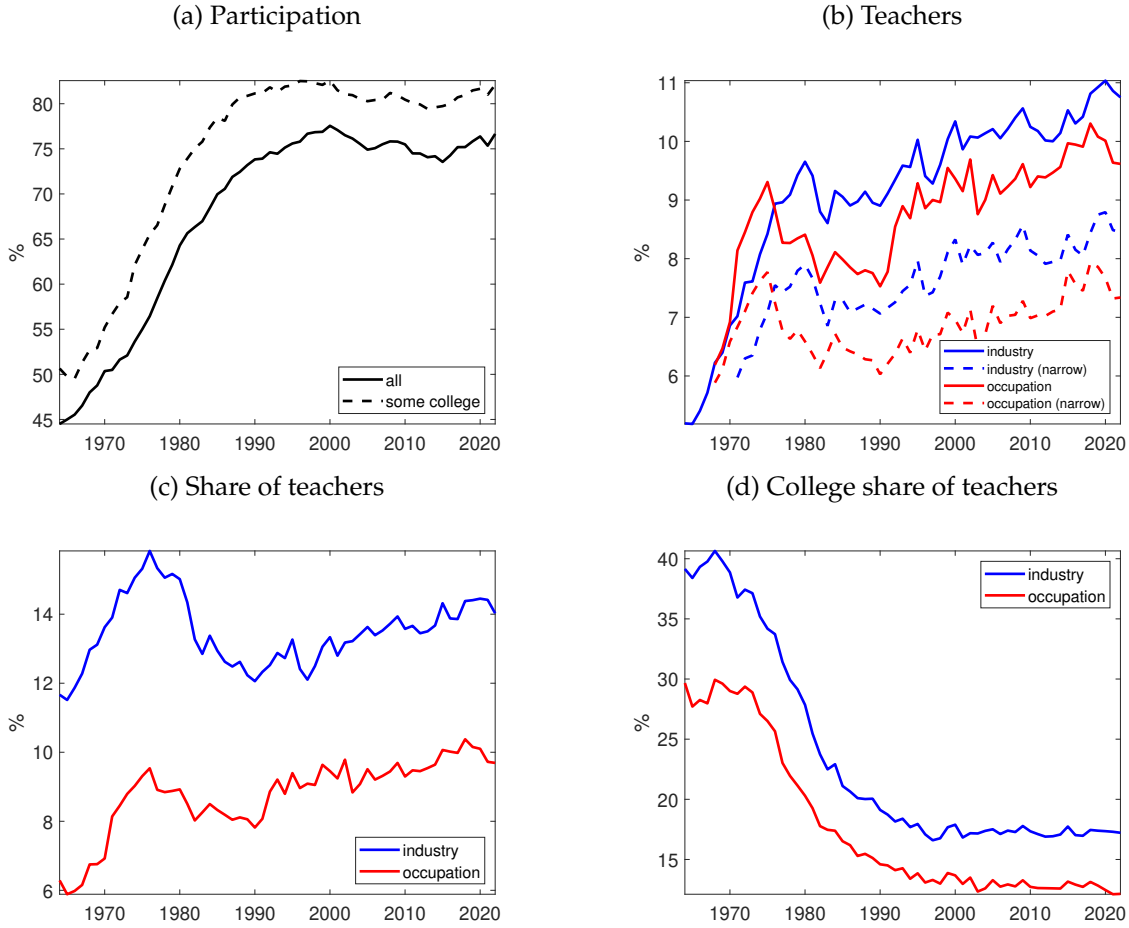
2 Empirical motivation

To motivate the model we present data from Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC). This is available on a consistent basis since 1964. We restrict the sample to working-age individuals aged between 25 and 55. We start by documenting the well-known increase in female labor force participation (Figure 2, Panel a). Between 1964 and 2000, female labor force participation increased from 44.5 percent to 77.6 percent. In the last 20 years, it has remained around this level, below the male participation rate (87.9 percent). Restricting attention to females with some college education a similar pattern is present, with the level shifted up by about 5 percentage points.

While the increase in female participation is well understood the reallocation across sectors is less well known. We study two definitions of "teachers" broadly covering the education sector. The first is based on industry SIC codes, the second based on an occupation definition (OCC1990).[3] Qualitatively, both definitions show the same trends, but the industry definition is broader and less skilled. Panel b plots the share of working-age

[3]Industry definition: Elementary and secondary schools (842), Colleges and universities (850), Vocational schools (851), Libraries (852), Educational services n.e.c. (860). Occupation definition: Postsecondary Teachers (2200) Preschool and Kindergarten Teachers (2300) Elementary and Middle School Teachers (2310) Secondary School Teachers (2320) Special Education Teachers (2330) Other Teachers and Instructors (2340) Teacher Assistants (2540) Education, Training, and Library Workers, nec (2550)

Figure 2: Female teachers in the labor force



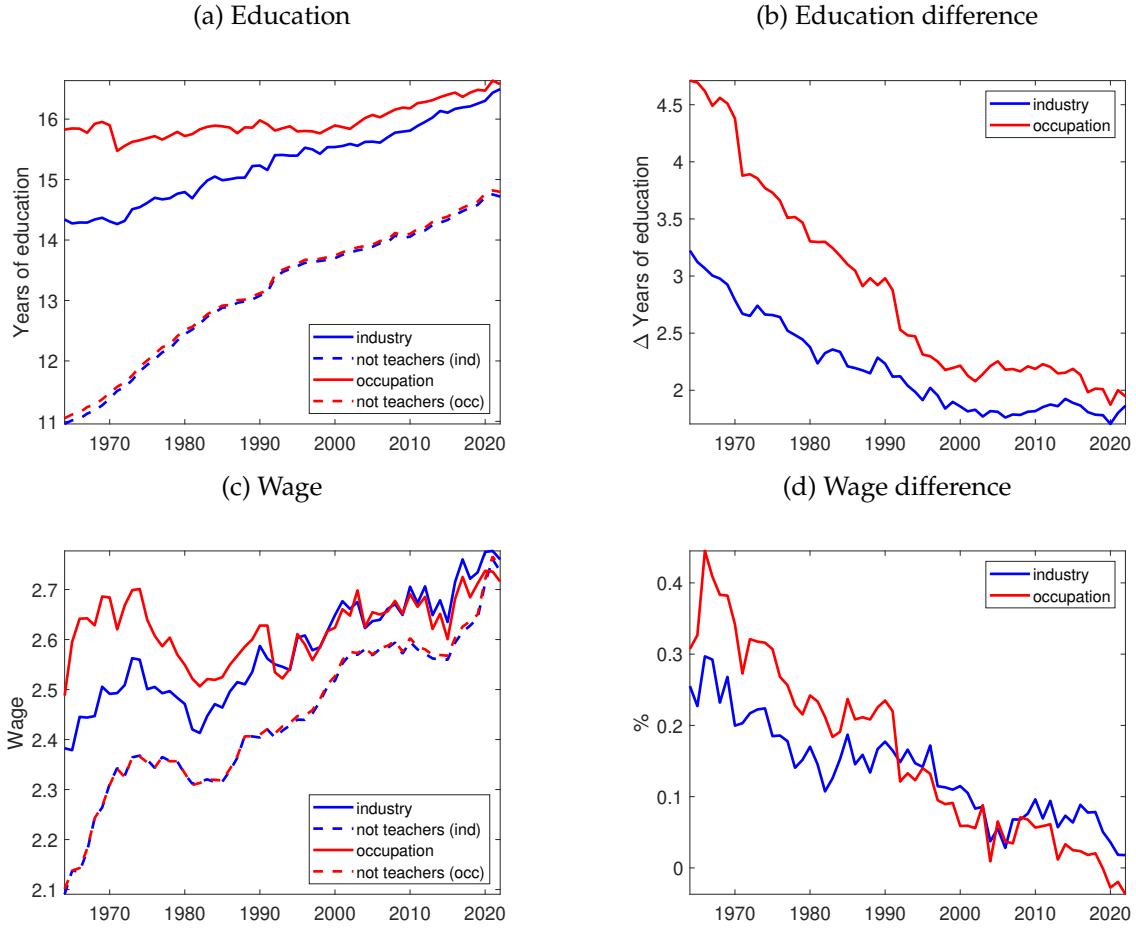
Notes: Data from the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC)

females that are teachers. In 1964, 5.2 percent of females were teachers. This increased quickly as women entered the labor force, reaching 9 percent in 1980. After this the share continued to increase at a slower rate, reaching 11 percent in 2020. Also shown is a narrower definition of teachers which focuses on "school teachers." For this set of workers, the level is lower and the increase is slightly less dramatic.

Panel c plots female teachers as a share of the female workforce. The long-run trend has been for this share to increase. It grew from 12 percent in 1964 (6 percent by occupation definition) to 14 percent in 2020 (10 percent by occupation definition). However, this masks medium-run dynamics. The share of females participating in the labor force teaching peaked in 1976 (industry definition). After this, the share of teachers fell as participation increased faster than the teaching workforce. The latter force is clearer if one only considers women with some college education.[4] This is a relevant definition as almost all teachers have some college education. In this group, the change is particularly striking. In 1968, 40.6 percent of college-educated women in the labor force were working as teachers (29.9 percent by occupation). By 2000, this had fallen to 17.9 percent and it has remained roughly

[4]Including those with one year of college but no degree.

Figure 3: Education and wages



Notes: Data from the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC)

at this level since (12.3 percent by occupation). This represents a dramatic reallocation of college-educated women in the labor force out of teaching and into the private market.

This reallocation can also be understood by looking at measures of the skill composition of the teaching workforce. Figure 3 plots the education level and wages of female workers in teaching and non-teaching jobs. Panel a shows years of education. Over time the education level of females has been increasing. Female teachers are on average more educated than their non-teaching counterparts. However, the average years of education outside of teaching has been rising much more quickly. As a result, relative to those in the private sector the years of education has been falling. This can be seen in Panel b. In 1964, the average teacher had 4.5 more years of education than the average market worker (occupation definition). By 2000 this had fallen to 2 years. The decline using the industry definition is slightly less. A similar and related trend is seen in average wages. Panel c shows log wages, this is labor income divided by hours. Wages are deflated by the CPI. Female wages have increased, both for teachers and market sector workers. For teachers, the increase has been about 20 log points, whereas for non-teachers the increase has been closer to 70 log points (occupation definition). As a result, the teacher wage

difference has fallen substantially. In 1968 the average female teacher earned 30 percent more than a female worker in the private sector. As of 2020, the difference had disappeared. This change reflects both changes in the relative education level of the two groups and differences in the working hours penalty. In the Appendix, we decompose this change in detail, but the bottom line is that the relative real wage of female teachers has fallen.

2.1 Cross-state correlations

We can also study the correlation between the changes in the labor market and educational outputs in the cross-section. To do this we combine demographic and occupational variables from the American Community Survey and CPS with educational attainment from the National Assessment for Educational Progess (NAEP). This provides a representative measure of academic achievement across US states and is available for a variety of grades and subjects. We pool data on maths and reading for grades 4, 8 and 12.

Our main specification takes the form:

$$\Delta G_j = \alpha + \beta \Delta FLP_j + \gamma \mathbf{X}_j + \delta(s, g) + \epsilon_j \tag{1}$$

where we regress the state j level change in the grade G_j between 1990 and 2022 (or the longest time horizon available) on the change in female labor force participation FLP_j between 1980 and 2022, in the ACS.[5] We also control for the change in the male participation rate, average income, education, the school age to working age population, shares by ethnic group $\mathbf{X}_j$, and subject and grade fixed effects $\delta(s, g)$

The results are presented in Table 1. Columns 1-3 show the specifications for female participation. The coefficient is significant, economically meaningful and importantly has the sign consistent with our theory. States that have seen a larger increase in the female participation rate have seen a bigger decline in educational attainment. Of course, this regression does not have a causal interpretation. The addition of controls or normalizing the grades lowers the coefficient a little. The magnitude is that 10 percentage point increase in female participation is associated with a 2.5 percent decline in average grades. The standard deviation of log changes in grades at the state level is 3.9 percent.

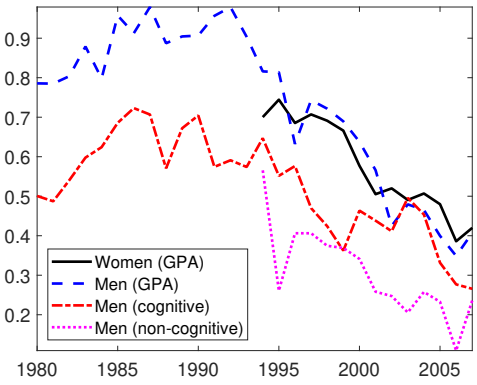
We also proxy for state variation in changes in the teacher quality. We use the average years of education of females in the teaching sector, defined using the broad industry-based definition. Here we observe a positive relationship across states, where the change in average educational level of female teachers is positively associated with better educational outcomes. This correlation is consistent with some link between the quality of educational inputs and educational outputs.

[5]The ACS allows us to look at the state-level change as it features a larger sample. This is particularly important for the education of teachers variables.

Table 1: Cross-state correlations

	Female participation			Teacher quality		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔFLP	-0.28	-0.25	-0.25		-0.30	-0.29
	(0.08)	(0.09)	(0.09)		(0.10)	(0.09)
$\Delta FT\ YEdu$				0.00	0.04	0.03
				(0.02)	(0.02)	(0.02)
ΔMLP		0.37	0.34		0.29	0.28
		(0.25)	(0.24)		(0.26)	(0.25)
Controls		✓	✓		✓	✓
Normalize grade			✓			✓
N	256	256	256	256	256	256
R^2	0.07	0.55	0.40	0.00	0.57	0.42

2.2 Swedish administrative data

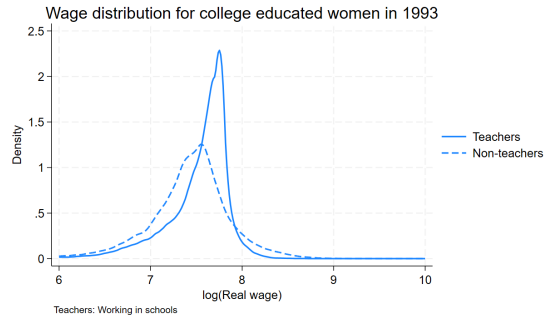


We also examine the cross-sectional patterns in the Swedish administrative data. We begin with the motivating evidence from Grönqvist and Vlachos (2016) which documents a decline in teacher quality in Sweden since at least the 1990s. The authors document a fall in the military grades of male teachers, during this period. The military grades should be a good guide of fundamental IQ and have been utilized in many studies. These are unavailable for females, but the authors show a.) GPA appears to move closely with military test scores, b.) the GPA scores of female teachers has fallen during the time period the data is available. This is consistent with falling teaching input quality. In the US this relationship has shown to be more prevalent for women than men, consistent with the CPS years of education presented above Bacolod (2007).

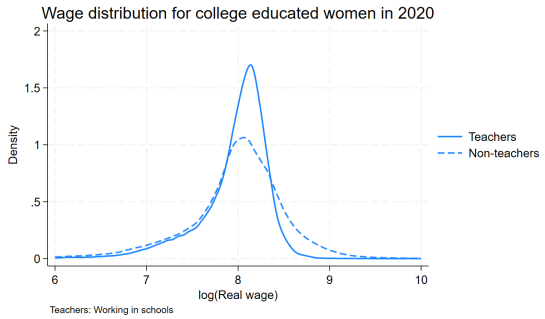
The distribution of income is also potentially informative for changes in the skill distributions across sectors. We compare the distribution of log income for teachers and non-teachers. Unsurprisingly, the teacher income distribution is more compressed than the non-teachers distribution.[6] Strikingly, there has been a clear shift in the non-teachers income distribution between 1993 and 2020. In 1993, there was a significant mass of private

[6]For teachers we use the definition: working in schools.

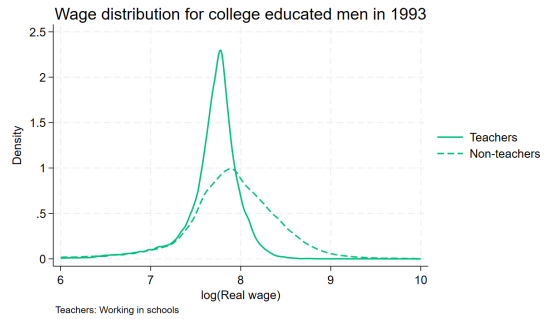
(a) women (1993)



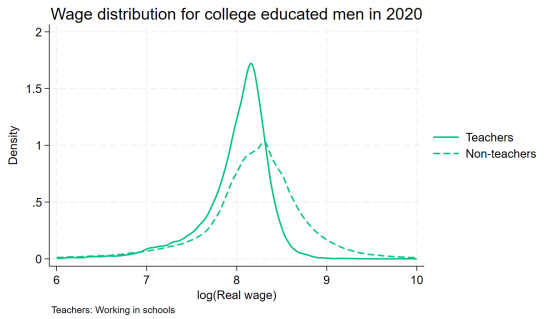
(b) women (2020)



(c) men (1993)



(d) men (2020)



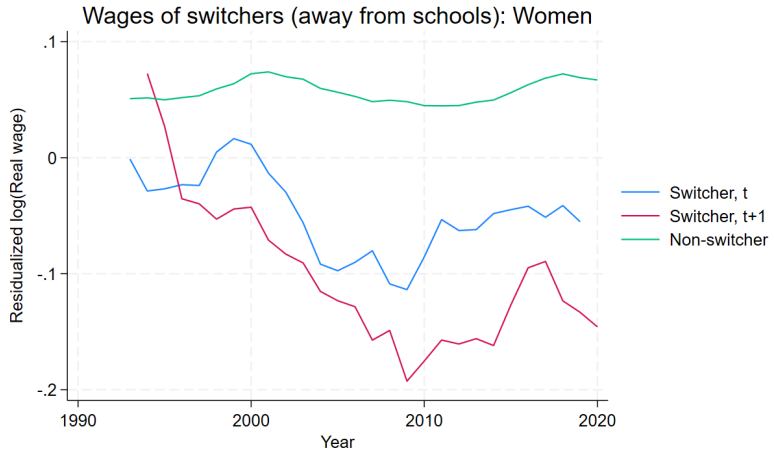
sector wages to the left of the teacher wage distribution. By 2020 the excess mass had moved to the right of the teacher wage distribution. This is consistent with two features of our theory: 1.) a fall in a private sector wage gap facing females, or 2.) a change in the relative distribution of skills in the private sector for females. Importantly, these patterns are not seen in the male income distributions over the same period, so it is not driven by the relative performance of the sectors. It is also not purely a feature of the data in the particular years, the shift can be observed to have taken place over time.

We can also look at the income of those on the margin of being teachers or private sector workers. In particular, we look at the residualized wages of switchers: those that were teachers and moved to the private market. This should provide information on the skill level of teachers in the private market. To do this we residualize wages by private sector and teaching sector, each year over time. We then take the average of the wage females that were teachers in period (t) and non-teachers in period (t+1). This is shown by black line in Figure 6. This has fallen substantially since 1993, indicating that females on the margin of being teachers and private sector working early were higher skilled (if wages are a reliable measure of productivity). We also observe switchers wages in the teaching profession prior to switching have fallen slightly, although not to the same extent.

3 Model

In this section we build a simple static model to study the allocation of women with different talents into three types of activities: (H)ome production, (T)eaching, and (M)arket

Figure 6: Decline in wage of switchers



work. We will compare the optimal allocation chosen by the social planner to the private allocation determined by the market. The allocation determined by the market will be affected by the two policy instruments available to the government: the teacher wage and the threshold ability requirement to become a teacher. The government will use these policy instrument so to minimize the misallocation of talents in the economy.

To focus on the allocation of women across the three activities, we assume that men work full time in the market. The talent for working at home is the same for all women, and denoted a^H. The talent for a woman working as a teacher, or in the market, is denoted, a^T and a^M, respectively, where these talents are distributed according to some multivariate distribution $f(a^T, a^M)$, with the corresponding CDF denoted $F(a^T, a^M)$.

Human capital For simplicity, we assume that the human capital of women are given. However, the allocation of women into teaching will affect the human capital of men, and thus determines how productive men are on the market. More specifically, the production of human capital in the economy is captured by the following decreasing returns to scale production function:

$$z = A_T N_T^{\alpha} \tag{2}$$

where A_T is a positive productivity parameter and where $\alpha \leq 1$. We consider decreasing returns to capture some fixed factor in educational production, such as the school capital stock. The total effective labor supply of teachers, N_T, is also central for the human capital in the economy, and is given by

$$N_T = \iint a^T \mathcal{T}(a^T, a^M) F(da^T, da^M) \tag{3}$$

where $\mathcal{T}(a^T, a^M)$ is an indicator function for becoming a teacher, and takes on value 1 if a woman with abilities (a^T, a^M) is allocated to teaching.[7]

Production Total production is given by $Y = Y_M + Y_H$, where market output is produced by both men and women, i.e., $Y_M = A_M(N_M^m + N_M^w)$, where $A_M > 0$, whereas home prouction, Y_H, is produced by women only. For simplicity, we take capital and other inputs as given, and assume linear production functions in labor for both home and market production.

Using the effective labor supply of men in market work, $N_M^m = \int\int z a^M F(da^T, da^M) = Sz\mu^M$, where μ^M is the unconditional mean of a^M, and of women in market work, $N_M^w = \int\int a^M(1-\tau)\mathcal{M}(a^T, a^M)F(da^T, da^M)$, we can write market production as:

$$Y_M = \int\int a^M(1-\tau)\mathcal{M}(a^T, a^M)F(da^T, da^M) + z\mu^M, \tag{4}$$

$$Y_H = \int\int \left(1 - \mathcal{M}(a^T, a^M) - \mathcal{T}(a^T, a^M)\right) F(da^T, da^M), \tag{5}$$

$$z = A_T \left(\int\int a^T \mathcal{T}(a^T, a^M)F(da^T, da^M)\right)^{\alpha}. \tag{6}$$

where $\mathcal{M}(a^T, a^M)$ is an indicator function that takes on value 1 if a woman with abilities (a^T, a^M) is allocated to market work. The term $1-\tau$, with $\tau > 0$, captures that women face an output loss when working in the market sector. This wedge can in a broad sense represent that women have been discriminated on the labor market.

If we normalize the price on consumption to unity, we have the market wage per effective work hour given by $w_M = A_M$.

3.1 Socially optimal allocation

Here we consider the social planner's allocation of women across home, teaching, and market work. The planner can't change the wedge τ, but can allocate women optimally across sectors besides that. With an interior solution, i.e. one where the planner allocates some women to all sectors, the social planner's allocation is given by:

$$\mathcal{T}(a^T, a^M) = \begin{cases} 1 & a^T \geq \frac{a^M(1-\tau)A_M}{A_T\alpha\mu^M N_T^{\alpha-1}} \;\&\; a^T \geq \frac{a^H}{A_T\alpha\mu^M N_T^{\alpha-1}} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

[7] Alternatively, instead of viewing $\mathcal{T}(a^T, a^M)$, and later $\mathcal{M}(a^T, a^M)$ and $\mathcal{H}(a^T, a^M)$, as indicator functions, view them as a continues measure of work hours with total hours for each woman normalized to unity, $\mathcal{T}(a^T, a^M) + \mathcal{M}(a^T, a^M) + \mathcal{H}(a^T, a^M) = 1$. The planner, or women, can allocated the total time to the three different activities, T, M and H. The corner solutions in the allocation of hours are endogenous outcomes.

$$\mathcal{M}(a^T, a^M) = \begin{cases} 1 & a^M(1-\tau) > \frac{A_T \alpha \mu^M N_T^{\alpha-1} a^T}{A_M} \ \& \ a^M(1-\tau) \geq \frac{a^H}{A_M} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

and the effective stock of teachers are pinned down by the following fixed point equation:

$$N_T = \int_{\frac{a^H}{A_T \alpha \mu^M N_T^{\alpha-1}}}^{\infty} \int_0^{\frac{A_T \alpha \mu^M N_T^{\alpha-1} a^T}{A_M(1-\tau)}} a^T f(a^T, a^M) da^M da^T = g(N_T) \tag{9}$$

3.2 Optimal constrained policy

In reality in most countries the contract faced by teachers is restricted and often does not fully reflect the skill level a^T of teachers. To incorporate this friction we consider a setting where the social planner is constrained to offer a simple contract offered to all teachers. The teacher contract is composed of two instruments: a wage w and minimum required skill level a^T_{MIN}. All teachers receive the same wage regardless of their productivity at teaching or their outside option. There is no bargaining.

Women make an occupation decision given their productivity (a^T, a^M) and the offered contract (w, a^T_{MIN}). The decision problem is as follows, for a woman with (a^T, a^M):

$$\max_{\mathcal{M},\mathcal{T}} \left\{ a^H, w \mathbf{1}_{a^T \geq a^T_{MIN}}, A_M a^M (1-\tau) \right\}$$

where the first term captures home production, the second term working as a teacher and the third term working in the market.

If $a^T < a^T_{MIN}$, a female worker cannot choose to be a teacher. They choose to work in the market $\mathcal{M} = 1$ if $\tilde{a} > \frac{a^H}{A_M}$. Alternatively, if $a^T \geq a^T_{MIN}$, the worker first considers choice between $\mathcal{T}$ and $\mathcal{M}$. They prefer $\mathcal{M}$ over $\mathcal{T}$ if $(1-\tau)a^M > \frac{w_T}{A_M}$ otherwise they prefer to be a teacher. Then they compare working with home production. They chooses to work at home if $a^H > \max\left\{A_M a^M(1-\tau), w\right\}$. Notice if $w < a^H$ the equilibrium would feature no teachers.

3.3 Decentralized planner's problem

Given the household choice problem the planner chooses the policy instruments, (w, a^T_{MIN}), to maximize welfare. This means solving the following planning problem, where the

household choices have been incorporated into the definitions of Y_M, Y_H and N_T.

$$\max_{a^T_{MIN}, w} \quad Y_M + Y_H$$

where

$$\begin{aligned}
Y_M &= A_M(1-\tau)\int_0^{a^T_{MIN}} \int_{\frac{a^H}{a^M(1-\tau)}}^{\infty} a^M f(a^T, a^M) da^M da^T \\
&\quad + A_M(1-\tau)\int_{a^T_{MIN}}^{\infty} \int_{\frac{w^T}{A_M(1-\tau)}}^{\infty} a^M f(a^M, a^T) da^M da^T \\
&\quad + A_M \mu^M A_T\left(N_T\right) \\
Y_H &= a^H \int_0^{a^T_{MIN}} \int_0^{\frac{a^H}{A_M(1-\tau)}} f(a^M, a^T) da^M da^T \\
N_T &= \int_{a^T_{MIN}}^{\infty} \int_0^{\frac{w}{A_M(1-\tau)}} a^T f(a^T, a^M) da^M da^T
\end{aligned}$$

The solution to the planner's problem is characterized by the following two first order equations. Firstly, for the optimal minimum skill level.

$$\underbrace{A_M(1-\tau)\int_{\frac{a^H}{A^M(1-\tau)}}^{\frac{w}{A_M(1-\tau)}} a^M F(a^M, a^T_{MIN}) da^M}_{Y_M \uparrow \text{ as women choose M when not allowed in T when } a^T_{MIN} \uparrow}$$

$$\underbrace{-\alpha A_M A_T \mu^M {N_T}^{\alpha-1} \int_0^{\frac{w}{A_M(1-\tau)}} a^T_{MIN} F(a^M, a^T_{MIN}) da^M}_{Y_M \downarrow \text{ as men have less HC in M when } a^T_{MIN} \uparrow}$$

$$\underbrace{+a^H \int_0^{\frac{a^H}{A_M(1-\tau)}} F(a^M, a^T_{MIN}) da^M}_{Y_H \uparrow \text{ as women chose H when not allowed in T when } a^T_{MIN} \uparrow} \qquad = 0$$

The first line captures the gain from increasing a^T_{MIN}, by reallocting those that prefer to work in the market rather than work at home to produce market goods. The second line captures the loss, whereby human capital fall as a^T_{MIN} is increased because the effective units of the teacher input into human capital production falls. The third line captures the gain due to those no longer teaching moving back into home production which produces a^H. Note that a^T is fixed at $a^T = a^T_{MIN}$ in the distribution, when integrating over a^M in $F(a^M, a^T_{MIN})$.

Secondly, is the condition determing the optimal teacher wage.

$$\underbrace{-w\int_{a^T_{MIN}}^{\infty} F(\frac{w}{A_M(1-\tau)}, a^T)da^T}_{Y_M \downarrow \text{ as women choose T instead of M when } w\uparrow} + \underbrace{\alpha A_M A_T \mu^M {N_T}^{\alpha-1}\int_{a^T_{MIN}}^{\infty} a^T F(\frac{w}{A_M(1-\tau)}, a^T)da^T}_{Y_M \uparrow \text{ as men have more HC in M as more women choose T when } w\uparrow} = 0$$

The first line captures the loss of output from worker reallocating from the market sector to teaching. By indifference, the productivity of these workers in the market sector equals the teachers wage, $w = A_M a^M(1-\tau)$. The second line capture the marginal gain from increasing the teacher input into human capital production. This can also be written more compactly as:

$$\int_{a^T_{MIN}}^{\infty} (\alpha A_M A_T \mu^M {N_T}^{\alpha-1} a^T - w) f(\frac{w}{A_M(1-\tau)}, a^T)da^T = 0$$

Note that a^M is fixed at $a^M = \frac{w}{A_M(1-\tau)}$ in the distribution, when integrating over a^T in the FOC.

Finally, observe that the barrier to market production, τ, features in both the first order conditions. As the productivity barriers facing females decline, it is optimal for the social planner to choose a new optimal contract for teaching. This is despite the fact that in this simple case human capital production does not increase female workers productivity in the market sector.

4 Perfectly correlated ability: $a^T = a^M = a$

In this section we compare the most efficient allocation(s) of talents feasible to a government with access to two policy instruments to the optimal unconstrained allocation(s) a social planner would choose. The two policy instruments available to the government are the wage level of teachers, w, and the threshold ability level to become a teacher, a^T_{MIN}.

As the focus of the analysis is on discrimination, we assume that women and men, on average, are equally productive in market work in absence of discrimination (μ is the unconditional mean of men's and women's ability in market work, and we normalize this mean to unity for ease of presentation). Also, for analogous reasons, we assume that the productivity scaling parameter in the production functions is the same for all types

of work and normalized to unity, i.e., $A^H = A^M = A^T = 1$. It is straight forward to relax these assumptions, and it will not affect the generality of the results.

We will impose some structure on the distribution of talents or, as we often use, abilities. First, we make the following assumptions about the distribution of abilities:

Assumption 1. *The ability to work on the market or as a teacher is the same, $a^T = a^M = a$, where $F(a)$, $a \in R_+$, represents the cumulative distribution of a.*

Next we will derive the socially optimal allocation(s) of talents, as well as how the policy instruments can be used to reach the socially optimal allocation(s), both with and without discrimination. To provide a better intuition for the model, we will discuss the results both when the human capital is produced under constant and decreasing returns to teacher input.

4.1 Human capital production with constant returns to teaching ($\alpha = 1$) and no discrimination ($\tau = 0$)

The results when there is no discrimination against women in market work, $\tau = 0$, and human capital is produced under constant returns to teaching, $\alpha = 1$, is presented below.[8]

Result 2. *Given Assumption 1, $\tau = 0$, and $\alpha = 1$:*

1. *the socially optimal allocations of talents are such that women with $a \leq a^H$ will work at home, and women with $a > a^H$ will either work on the market or as teachers, thus $F(a^H)$ women in home production and $1 - F(a^H)$ women work on the market or in teaching,*

2. *the socially optimal allocations of talents can be obtained by policies (a^T_{MIN}, w) that fulfill: $a^T_{MIN} \in [a^H, \bar{a}]$, where $\bar{a} \in R_+$ is the highest ability in the distribution, and $w \in [0, \bar{a}]$ or $a^T_{MIN} \in [0, a^H)$ and $w \in [0, a^H)$.*

The intuition is as follows. First, to allocate women with low ability (women with $a \leq a^H$) to home production is key to obtain the socially optimal allocation when there is no discrimination and human capital is produced under constant returns to teaching. To allocate women to teaching or market work, on the other hand, leaves total production unaffected in this case. When allocated into teaching, women facilitate production by providing men with human capital to produce, whereas when allocated to market work, women produce the output themselves, and they contribute equally much to production in both these activities.[9]. Thus the social planner is indifferent between allocating women to

[8] if indifferent between H and T choose H, if indifferent between H and M, choose H, if indifferent between T and M, choose T or M (strange not going for strict in the last case, but for a reason that is not necessarily good). Also, as there are many allocations that are possible, one can in itself derive different allocations that are interchangeable by just shifting what to do in the indifference cases. Those distinctions are not considered here.

[9] this holds as we have made a number of assumption: men and women are equally productive absent discrimination, $\mu = 1$ and the same for women, $A^i = 1, i = H, M, T$, same mass of women and men, etc.

market work and to human capital production, which implies that there is a large number of potential allocations of women that are socially optimal.

The minimum threshold to be able to become a teacher and the teacher wage can be combined in a number of ways so to reach a socially optimal outcome. For example, preventing low ability women from becoming teachers by setting the threshold level, a^T_{MIN}, at least as high as the ability to produce at home, a^H, assures that low ability women chose to allocate themselves to home production. Their alternative is to choose market work, but the wage facing these low ability women on the market is too low for that to be an attractive option. The teacher wage plays no role for total production in this case as the choice to become a teacher or work in the market for women has no affect on production. It is only if the threshold level, a^T_{MIN}, is set below a^H, that the teacher wage is needed as a complement, $w < a^H$, to reduce the incentives for low ability workers to become teachers instead of working at home.

Notice that it is not important how the mass is allocated in the distribution for the results. There may, for example, be the case that no women have an ability such that $a > a^H$. In this case, simply no women should work at home, i.e., $F(a^H) = 0$ in the socially optimal allocation. Analogously, if no women have an ability such that $a > a^H$, all women should work at home. These women are then more productive in home production than if working on the market or providing men with human capital for market work through teaching.

Below are examples of socially optimal feasible allocations when the government has access to the policy instruments, a^T_{MIN} and w. For ease of presentation, the upper ability level in the distribution is denoted $\bar{a} \in R_+$, and the policy instrument a^T_{MIN} and w are never set above the upper ability.

- if $a^H < a^T_{MIN} < \bar{a}$, women with $a \in (a^H, a^T_{MIN})$ will work in the market as they cannot become teachers irrespective of w, women with $a \in [a^T_{MIN}, \bar{a}]$ will become teachers if $a \leq w$ and work on the market if $a > w$, thus $F(a^H)$ women in home production, and $F(w) - F(a^T_{MIN})$ in teaching, and $1 - F(w) + F(a^T_{MIN}) - F(a^H)$ women in market work, (see Figure 7 Panel a),

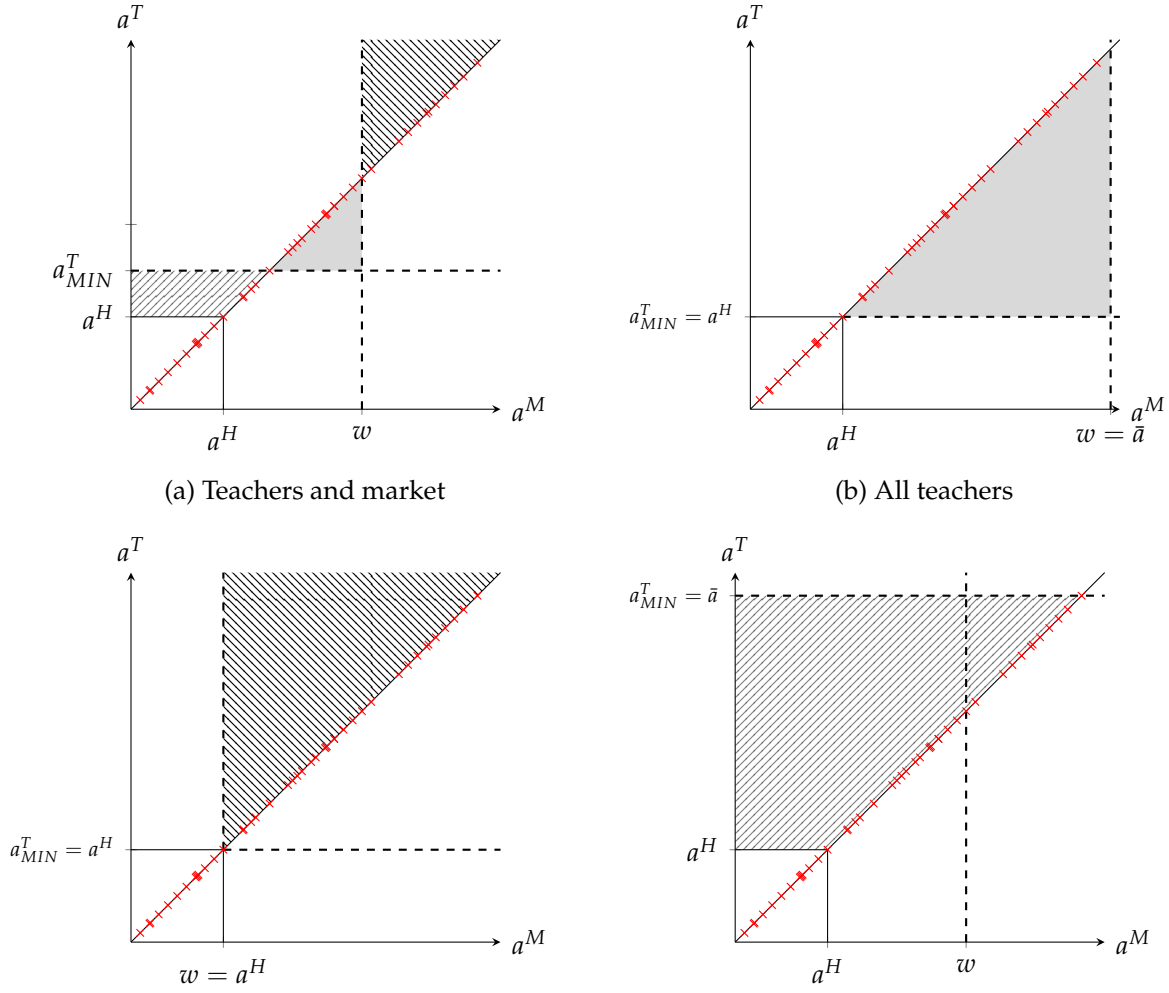
- if $a^T_{MIN} = a^H$, women with $a > a^H$ will become teachers if $a \leq w$ and work on the market if $a > w$, thus $F(a^H)$ women in home production, and $F(w) - F(a^H)$ in teaching, and $1 - F(w)$ women in market work, (see Figure 7), Panels b and c for the case when $w = \bar{a}$ and $w = a^H$, respectively),

- if $a^T_{MIN} = \bar{a}$, no women will become teachers irrespective of w,[10] thus $F(a^H)$ women in home production and $1 - F(a^H)$ women in market work (Panel d)

[10]This statement builds on that if a woman is indifferent between being in M or T, she will choose T. That is, if $w = \bar{a}$, women with ability $\bar{a}$ are, in fact, indifferent between becoming teachers or work in the market. Here, we assume that they work in the market.

- if $a^T_{MIN} < a^H$ and $w < a^H$, no women would want to become a teacher, thus $F(a^H)$ women in home production, and $1 - F(a^H)$ women in market work.

Figure 7: Possible allocations with constant returns to teaching



(a) Teachers and market

(b) All teachers

(c) All market

(d) All below a^T_{MIN}

Notes: in Panel a: policy: $w \in (a^H, \bar{a})$ and $a^T_{MIN} \in (a^H, \bar{a})$ Women in H, M, T, and M; in Panel b: policy: $a^T_{MIN} = a^H$ and $w = \bar{a}$ Women in H and T; In Panel c: policy: $a^T_{MIN} = a^H$ and $w = a^H$ Women in H and M; in Panel d: policy: $w \in (a^H, \bar{a})$ and $a^T_{MIN} = \bar{a}$ Women in H and M. In all cases: $\alpha = 1$, $\tau = 0$.

4.2 Human capital production with constant returns to teaching ($\alpha = 1$) and discrimination ($\tau > 0$)

The results when there is discrimination against women in market work, $\tau > 0$, and human capital is produced under constant returns to teaching, $\alpha = 1$, is presented below.

Result 3. *Given Assumption 1 and $\tau > 0$, $\alpha = 1$:*

1. *the socially optimal allocation of talents are such that women with $a \leq a^H$ will work at home, and women with $a > a^H$ will be teachers, thus $F(a^H)$ women in home production and $1 - F(a^H)$ women in teaching,*

2. *the socially optimal allocation of talents can be obtained by a policy that fulfills:* $a^T_{MIN} = a^H$ *and* $w = \bar{a}$.

The intuition is as follows. First, when there is discrimination, and human capital is produced under constant returns to teaching, the socially optimal allocation is such that low ability women (women with $a < a^H$) would be allocated to home production whereas all women with high ability (women with $a < a^H$) should be teachers. As women and men are equally productive in market work in absence of discrimination, the presence of discrimination is going to make women relatively less productive in market work. It is therefor better to allocate women to teaching in order for them to provide men with the human capital needed for market production.

It is straightforward to obtain the socially optimal allocation with the available policy instruments. The government can attract women to teaching by setting a very high teacher wage, while at the same time prevent women with the lowest ability from becoming teachers instead of working at home, by setting the threshold level, a^T_{MIN}, equal to the ability to produce at home, a^H. See Figure 8 for this case.

Notice that it is not important how the mass is allocated in the distribution for the results. There may, for example, be the case that no women have an ability such that $a > a^H$. In this case, simply no women should work at home, i.e., $F(a^H) = 0$ in the socially optimal allocation. Analogously, if no women have an ability such that $a > a^H$, all women should work at home. These women are then more productive in home production than if providing men with human capital for market work through teaching.

4.3 Human capital production with decreasing returns to teaching ($\alpha < 1$) and no discrimination ($\tau = 0$)

When human capital is produced under decreasing returns to effective teacher input, $\alpha < 1$, the socially optimal allocation of women depends crucially on the size of the stock of effective teachers. We define the set(s) of abilities higher than a^H, which when women with these abilities are put into teaching, will generate the optimal stock of effective teachers, $N^*_T \mid_{\tau=0}$.

Definition 4. *Let* $A_{>a^H} = \{a : a > a^H\}$ *and* $A^T_{>a^H} \subseteq A_{>a^H}$ *if* $\int a f(a) da = \alpha^{\frac{1}{1-\alpha}} = N^*_T \mid_{\tau=0}$.

We can conclude:

Result 5. *Given Assumption 1,* $\tau = 0$, $\alpha < 1$, *and* $A^T_{>a^H}$ *exists:*

1. *the socially optimal allocations of talents are such that women with* $a \leq a^H$ *will work at home, women with* $a \in A^T_{>a^H}$ *will be allocated to teaching, and women with* $a > a^H$ *and* $a \notin A^T_{>a^H}$ *will be allocated to market work.*

Thus, $F(a^H)$ women in home production, $\int_{a \in A^T_{>a^H}} f(a)da$ women in teaching, and $1 - F(a^H) - \int_{a \in A^T_{>a^H}} f(a)da$ women in market work.

2. *the socially optimal allocations of talents can be obtained by (a^T_{MIN}, w) that fulfill:*[11] *$a^T_{MIN} \in [a^H, \bar{a})$, $w \in (a^H, \bar{a}]$*[12]*, where $\bar{a} \in R_+$ is the highest ability level in the distribution, and $\int_{a^T_{MIN}}^{w} aF(a)da = \alpha^{\frac{1}{1-\alpha}}$.*

To derive a socially optimal allocation of talents, it is again central that low ability women, women with $a \leq a^H$, are allocated to home work. Women with high ability, women with $a > a^H$, should be allocated to market work and teaching so that for each of these women the marginal productivity in teaching and market work is the same. For a woman with ability a, the marginal productivity in teaching is $a\alpha(N^T)^{\alpha-1}$, whereas the marginal productivity in market work is a. These productivities are the same though if $N_T = N_T^* = \alpha^{\frac{1}{1-\alpha}}$, and this holds for all a. Thus it is optimal to allocate women into teaching until the optimal stock of effective teachers is reached, and a reallocation of a women from market work to teaching, or vise versa, will have no impact on aggregate production at that stage. This is shown in Figure 8 Panel b.

It is also the case that it does not matter if high or low ability women, among women with $a <^H$, is allocated to teaching. It is clearly the case that if high, rather than low, ability women, among these women, are allocated to teaching, less women need to be allocated to teaching. This will, however, not make it more efficient to choose these women as they, at the same time, are also more productive in market work. Thus, from a social planner's point of view it does not matter which women, among women with $a > a^H$, that are allocated to teaching.

From a government perspective with access to the two policy instruments, a^T_{MIN} and w, there are a number of choices which can generate a socially efficient allocation. It is simply a matter of choosing a^T_{MIN} and w so to prevent women with $a \leq a^H$ to leave work at home for teaching and, at the same time, encourage just enough women with $a > a^H$ to become teachers so that the effective stock of teachers, $N_T^* = \alpha^{\frac{1}{1-\alpha}} \mid_{\tau=0}$.

Below we provide three examples of policy choices that generate an optimal effective stock of teachers, while at the same time prevent low ability women to leave home production for teaching. With these policy instruments, the stock of effective teachers are obtained by attracting women with abilities in a consecutive order in the ability distribution. However, it could be either low ability women, high ability, or middle ability women, among women with $a > a^H$, that are attracted into teaching. Notice, however, that the optimal stock of

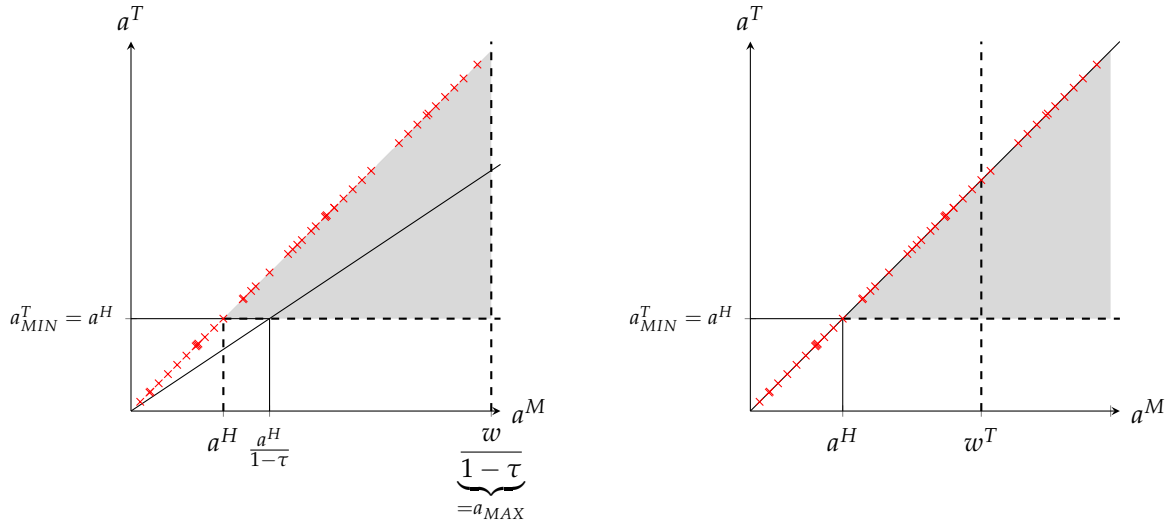
[11] If an individual is indifferent between home work and market work or teaching, he or she chooses home work, if an individual is indifferent between market work and teaching, he or she chooses teaching.

[12] The upper limit on the wage is never set above $\bar{a}$ as an assumption rather than anything else. In fact, the wage could be set above $\bar{a}$. If $a = \bar{a}$ and $w = \bar{a}$, payments to teachers are at least as high as for market work and all women, including those with the highest ability, would want to be teachers, considering that we assume that if one are indifferent between teaching and market work, one chooses teaching.

effective teachers can only be obtained if there exists a significant mass of women with their associated abilities in a way that makes this feasible (that is $a \in A^T_{>a^H}$ need to exist). It also clear that it is less likely that the optimal effective stock of teachers can be obtain by restricting the sample of women with $a > a^H$, by excluding women with the highest and lowest abilities in the teaching occupation.

- if $a^T_{MIN} = a^H$ and w chosen to fulfill $\int_{a^H}^{w} af(a)da = \alpha^{\frac{1}{1-\alpha}}$, women with $a \leq a^H$ in home, and women with $a \in (a^H, w]$ in teaching, and women with $a \in (w, \bar{a}]$ in market work, thus $F(a^H)$ women in home production, $F(w) - F(a^H)$ women in teaching, and $1 - F(w)$ women in market work,
- if a^T_{MIN} is chosen to fulfill $\int_{a^T_{MIN}}^{\bar{a}} af(a)da = \alpha^{\frac{1}{1-\alpha}}$ and $w = \bar{a}$, women with $a \leq a^H$ in home, and women with $a \geq a^T_{MIN}$ in teaching, and women with $a \in (a^H, a^T_{MIN})$ in market work, thus $F(a^H)$ women in home production, $1 - F(a^T_{MIN})$ women in teaching, and $F(a^T_{MIN}) - F(a^H)$ women in market work,
- if $a^T_{MIN} > a^H$ and $w < \bar{a}$ so to fulfill $\int_{a^T_{MIN}}^{w} af(a)da = \alpha^{\frac{1}{1-\alpha}}$, women with $a \leq a^H$ in home, and women with $a \in (a^H, a^T_{MIN})$ together with women with $a > w$ are in market work, and women with $a \in [a^T_{MIN}, w]$ in teaching, thus $F(a^H)$ women in home production, $F(w) - F(a^T_{MIN})$ women in teaching, and $1 - F(a^H) - F(w) + F(a^T_{MIN})$ women in market work,

Figure 8: Solution to the optimal human capital production



(a) Constant returns to teaching

(b) Decreasing returns to teaching

Notes: In Panel a: $\alpha = 1$, $\tau > 0$, Policy: $a^T_{MIN} = a^H$ and $\frac{w}{1-\tau} = \bar{a}$ Women in H and T; In Panel b: $\alpha < 1$, $\tau = 0$, Policy: w chosen so to get N^*, Women in H,T and M A $\alpha = 1$, $\tau = 0$ although a^T_{MIN} cannot be at $\bar{a}$.

4.4 Human capital production with decreasing returns to teaching ($\alpha < 1$) and discrimination ($\tau > 0$)

When human capital is produced under decreasing returns to effective teacher input, $\alpha < 1$, the socially optimal allocation of women depends crucially on the size of the stock of effective teachers. This holds also in the presence of discrimination. Define the set of abilities higher than $\frac{a^H}{1-\tau}$, which when women with these abilities are put into teaching, will generate the optimal stock of effective teachers, $N_T^* \mid_{\tau>0}$, in the presence of discrimination.

Definition 6. *Let* $A_{>\frac{a^H}{1-\tau}} = \{a : a > \frac{a^H}{1-\tau}\}$ *and* $A^T_{>\frac{a^H}{1-\tau}} \subseteq A_{>\frac{a^H}{1-\tau}}$ *if*

$$\int af(a)da = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}} = N_T^* \mid_{\tau>0}$$

We can conclude:

Result 7. *Given Assumption 1,* $\tau > 0$, $\alpha < 1$, *and* $A^T_{>\frac{a^H}{1-\tau}}$ *exists:*

1. *the socially optimal allocations of talents are such that women with* $a \leq \frac{a^H}{1-\tau}$ *will work at home, and women with* $a \in A^T_{>\frac{a^H}{1-\tau}}$ *will be allocated to teaching, and women with* $a > \frac{a^H}{1-\tau}$ *and* $a \notin A^T_{>\frac{a^H}{1-\tau}}$ *will be allocated to market work. Thus,* $F(\frac{a^H}{1-\tau})$ *women in home production,* $\int_{a\in A^T_{>\frac{a^H}{1-\tau}}} f(a)da$ *women in teaching, and* $1 - F(\frac{a^H}{1-\tau}) - \int_{a\in A^T_{>\frac{a^H}{1-\tau}}} f(a)da$ *women in market work.*

2. *the socially optimal allocations of talents can be obtained by* (a^T_{MIN}, w) *that fulfill:*[13] $a^T_{MIN} \in [\frac{a^H}{1-\tau}, \bar{a})$, $w \in (a^H, \bar{a}]$[14], *where* $\bar{a} \in R_+$ *is the highest ability level in the distribution, and* $\int_{a^T_{MIN}}^{\frac{w}{1-\tau}} af(a)da = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$.

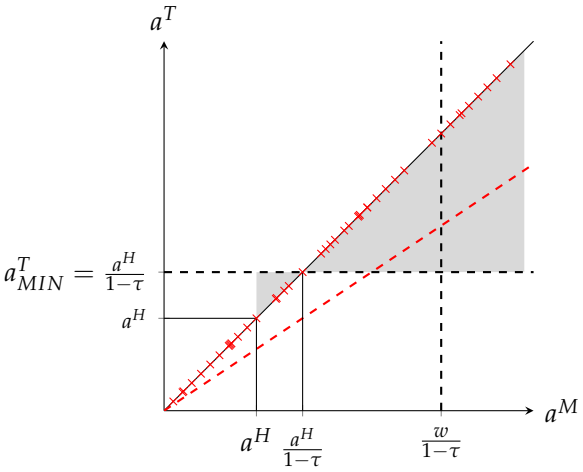
To derive a socially optimal allocation of talents, it is again central that low ability women are allocated to home work. In the presence of discrimination, $\tau > 0$, the threshold level for when women with low ability are equally productive in home production and market work has increased to $a = \frac{a^H}{1-\tau}$. This implies that more of women are allocated to home production. One would potentially think that women with $a \in [a^H, \frac{a^H}{1-\tau}]$ find it attractive to become teachers instead of working at home as women in teaching are not discriminated. However, this is not case due to adjustments in the effective stock of teachers. In fact we can conclude the following:

[13] If an individual is indifferent between home work and market work or teaching, he or she chooses home work, if an individual is indifferent between market work and teaching, he or she chooses teaching.

[14] The upper limit on the wage is never set above $\bar{a}$ as an assumption rather than anything else. In fact, the wage could be set above $\bar{a}$. If $a = \bar{a}$ and $w = \bar{a}$, payments to teachers are higher then for market work and all women, including those with the highest ability, would want to be teachers. Here, $w \geq \bar{a}(1-\tau)$ is enough to make all women choose teaching if they pass the minimum threshold level.

Result 8. *Given Assumption 1,* $\alpha < 1$*, and* $A^T_{>\frac{a^H}{1-\tau}}$ *exists*[15]*, the stock of effective teachers is larger in the presence of discrimination,* $N_T^* \mid_{\tau>0} = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}} > (\alpha)^{\frac{1}{1-\alpha}} = N_T^* \mid_{\tau=0}$.

Figure 9: Allocation with decreasing returns and discrimination



Notes: $\alpha < 1$, $\tau > 0$, Policy: w and a_{MIN} chosen so to get N^*, notice that the dotted red line is based on the stock of effective teacher when there are no discrimination (or any other stock of effective teachers that are too low).

As women are discriminated in market work relative to teaching, it is efficient to allocate more women into teaching. In fact, women will again be allocated into teaching until for each woman the marginal productivity in teaching is the same as the marginal productivity in market work. For a woman with ability a, the marginal productivity in teaching is $a\alpha(N^T)^{\alpha-1}$, whereas the marginal productivity in market work is $a(1-\tau)$. These productivities are the same though if $N_T = N_T^* \mid_{\tau>0} = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$, and this holds for all a. Thus it is optimal to choose home work instead of teaching for women with $a \in [a^H, \frac{a^H}{1-\tau}]$ as the marginal productivity in teaching is $a(1-\tau)$. This is shown graphically in Figure 15.

It is also the case in the presence of discrimination that it does not matter if high or low ability women, among women with $a > \frac{a^H}{1-\tau}$, is allocated to teaching. It is clearly the case that if high, rather than low, ability women, among these women, are allocated to teaching, less women need to be allocated to teaching. This will, however, not make it more efficient to choose these women as they, at the same time, are also more productive in market work. Thus, from a social planner's point of view it does not matter which women, of women with $a > \frac{a^H}{1-\tau}$, that are allocated to teaching. However, notice that the optimal stock of effective teachers can only be obtained if there exists a significant mass of women with their associated abilities in a way that makes this feasible (that is $a \in A^T_{>\frac{a^H}{1-\tau}}$ need to exist), and in the presence of discrimination this this is less likely to be the case.

[15] If $A^T_{>\frac{a^H}{1-\tau}}$ exists, then $A^T_{>a^H}$ exists as $A^T_{>\frac{a^H}{1-\tau}} \subseteq A^T_{>a^H}$. It is most likely that even if $A^T_{>\frac{a^H}{1-\tau}}$ does not exist and $A^T_{>a^H}$ exists, that is enough.

From a government perspective with access to the two policy instruments, a^T_{MIN} and w, there are a number of choices which can generate a socially efficient allocation. It is simply a matter of choosing a^T_{MIN} and w so to prevent women with $a \leq \frac{a^H}{1-\tau}$ to leave work at home for teaching and, at the same time, encourage just enough women with $a > \frac{a^H}{1-\tau}$ to become teachers so that the effective stock of teachers are given by $N_T = N^*_T = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$.

Below we provide three examples of policy choices that generate an optimal effective stock of teachers, while at the same time prevent low ability women to leave home production for teaching. With these policy instruments, the stock of effective teachers are obtained by attracting women with abilities in a consecutive order in the ability distribution. However, it could be either low ability women, high ability, or middle ability women, among women with $a > \frac{a^H}{1-\tau}$, that are attracted into teaching. Notice, however, that the optimal stock of effective teachers can only be obtained if there exists a significant mass of women with their associated abilities in a way that makes this feasible (that is $a \in A^T_{> \frac{a^H}{1-\tau}}$ exists). It also clear that it is less likely that the optimal effective stock of teachers can be obtain by restricting the sample of women with $a > a^H$, by excluding women with the highest and lowest abilities in the teaching occupation.

- if $a^T_{MIN} = \frac{a^H}{1-\tau}$ and w chosen to fulfill $\int_{\frac{a^H}{1-\tau}}^{\frac{w}{1-\tau}} aF(a)da = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$, women with $a \leq \frac{a^H}{1-\tau}$ in home, women with $a \in (\frac{a^H}{1-\tau}, \frac{w}{1-\tau}]$ in teaching, and women with $a \in (\frac{w}{1-\tau}, \bar{a}]$ in market work, thus $F(\frac{a^H}{1-\tau})$ women in home production, $F(\frac{w}{1-\tau}) - F(\frac{a^H}{1-\tau})$ women in teaching, and $1 - F(\frac{w}{1-\tau})$ women in market work,

- if a^T_{MIN} is chosen to fulfill $\int_{a^T_{MIN}}^{\bar{a}} aF(a)da = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$ and $w = \bar{a}$, women with $a \leq \frac{a^H}{1-\tau}$ in home, and women with $a \geq a^T_{MIN}$ in teaching, and women with $a \in (\frac{a^H}{1-\tau}, a^T_{MIN})$ in market work, thus $F(\frac{a^H}{1-\tau})$ women in home production, $1 - F(a^T_{MIN})$ women in teaching, and $F(a^T_{MIN}) - F(\frac{a^H}{1-\tau})$ women in market work,

- if $a^T_{MIN} > \frac{a^H}{1-\tau}$, $\frac{w}{1-\tau} < \bar{a}$, and fulfills $\int_{a^T_{MIN}}^{\frac{w}{1-\tau}} aF(a)da = (\frac{\alpha}{1-\tau})^{\frac{1}{1-\alpha}}$, women with $a \leq \frac{a^H}{1-\tau}$ in home, and women with $a \in (\frac{a^H}{1-\tau}, a^T_{MIN})$ together with women with $a > \frac{w}{1-\tau}$ are in market work, and women with $a \in [a^T_{MIN}, \frac{w}{1-\tau}]$ in teaching, thus $F(\frac{a^H}{1-\tau})$ women in home production, $F(\frac{w}{1-\tau}) - F(a^T_{MIN})$ women in teaching, and $1 - F(\frac{a^H}{1-\tau}) - F(\frac{w}{1-\tau}) + F(a^T_{MIN})$ women in market work,

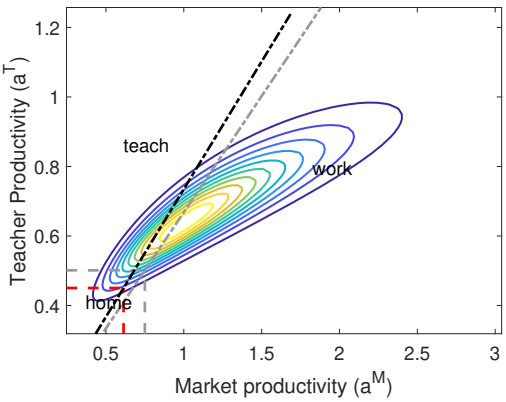
5 Predictions of the simple model

Before turning to the full quantitative exercise, we can build further intuition for the model's mechanism by examining how the allocation responds to a change in the labor market barrier, τ. We relax the assumption of the previous section and allow market production skill a^M and teaching skill a^T to be positively correlated, with correlation

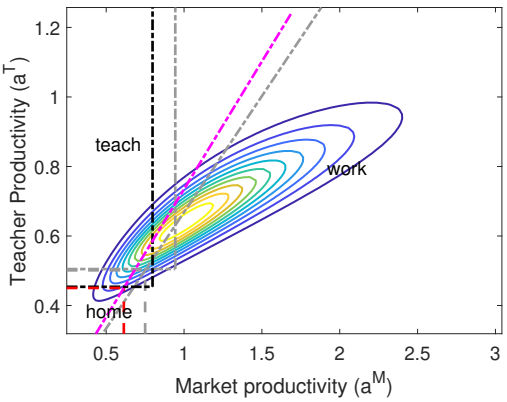
coefficient less than 1.[16] Figure 10 presents the allocations. The left panel presents the unconstrained allocation and the right panel presents the constrained teacher contract. The ellipses indicate where the mass of productivity types is located.

Figure 10: Characterization of allocation

(a) Unconstrained allocation



(b) Constrained allocation



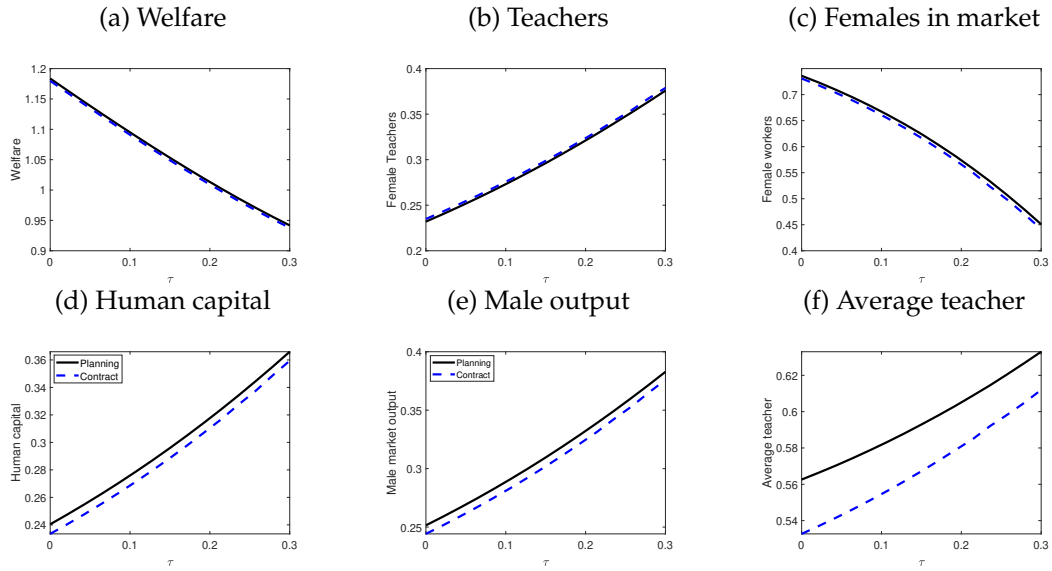
Notes:

Panel a shows the unconstrained allocation. The gray lines are for a high τ, while the black lines are for a low τ. Low teacher and market productivity females are assigned to home production, the bottom left rectangle. The diagonal line shows the planner's indifference defined by equation 7. Females with high teacher productivity and lower market productivity are assigned to teaching, females with high market productivity and relatively lower teacher productivity are assigned to the market. When the market barrier, τ, falls the indifference condition tilts anti-clockwise. Teachers with high market productivity are optimally allocated into the market. As the real market wage has increase females are also reallocated out of home production into both teaching and the market.

Panel b shows what happens under the constrained teacher contract. As the planner is only able to choose a teacher wage, w^T, and minimum skill, a^T_{min}, they are unable to replicate the first best. Instead, the planner chooses a wage and skill cut-off to implement the allocation as best as possible. The teacher allocation is instead the top-left rectangle. This implies two distortions. There are high skill teacher's the planner would like to allocate to teaching but instead are allocated to the market (top triangle). There are also low productivity teachers the planner would like to allocate to market but instead choose to be teachers due to the high teacher wage (triangle below the indifference curve). The constrained planner trades-off these two inefficiencies. When τ falls the constrained planner alters the contract to replicate the shifting indifference curve. The teacher wage and minimum skill level decline. The rectangle which denote the optimal constrained teacher allocation most down and to the left.

[16]We assume joint log normal distribution of these skills as in the full Quantitative model.

Figure 11: Changing the barrier in a simple model



Notes:

We can also see how important aggregate quantities respond to a change in the market barrier, τ. Figure 11 plots these against an illustrative market barrier between 0 and 0.3 on the x-axis. The black line shows the unconstrained planning allocation. The blue dashed line shows the constrained teacher contract. Panel a plots welfare. Welfare is strictly decreasing in the market barrier as it lowers total potential market production. We also see that in the simple model the distortion implied by the constrained teacher contract is fairly small. Indeed, the planner and contract allocation deliver similar welfare. For lower values of τ, the share of females allocated to teaching falls (Panel b), while the share allocated to the market increases (Panel c). This is the optimal response of the planner in the face of the higher productivity value of females in the marketplace. However, as a result human capital declines (Panel d), which generates a fall in male output (Panel e). What is important to realize is that the decline in effective units of teacher skill is entirely optimal. Higher-quality teachers are reallocated into the market sector as their relative market productivity increases. Under this condition, it is optimal to produce a smaller human capital input. The fall in male output is a side effect of this reallocation.

In the panels in the bottom row we see a larger deviation between the unconstrained allocation and optimal teacher contract. The contract makes less efficient use of the available teaching skill. Slightly, more teachers are employed (Panel b), but they generate less human capital (Panel d). Another closely related implication, is that the lower value of τ implies a lower average quality of teacher, defined as:

$$AveT = \frac{\int a^T \mathcal{T}(a^T, a^M) f(a^T, a^M) da^T da^M}{\int \mathcal{T}(a^T, a^M) f(a^T, a^M) da^T da^M} \tag{10}$$

as fewer high-skill teachers are substituted for more low-skill teachers. Further, we see that the constrained contract delivers a lower average teacher quality allocation, regardless of the level of τ.

6 Quantitative Model

To take the model to the data we require four extensions: 1.) a high-skill barrier, 2.) population trends, 3.) male teachers, 4.) complementarities in human capital

High-skill barrier: Firstly, we allow for two discriminatory barriers, which impact different parts of the distribution. In addition to the average wedge, τ, we add a high-skill wedge, $\hat{\tau}$. This takes the form of a power function $x^{1-\hat{\tau}}$ and implies higher productive women face a larger discriminatory barrier. It is motivated by the findings of Goldin (2014), who provide evidence of "greedy" jobs, still proving to be a barrier to work or equal compensation for high-skilled women. Now the productivity of women in the market is given by:

$$\tilde{a}^M = (1-\tau)(a^M)^{1-\hat{\tau}} \tag{11}$$

As with the average skill barrier, we consider changes in both barriers over time.

Population shares: Secondly, we allow for changing population shares. There are two important trends: 1.) the increase in college-educated workers seen in all developed countries, 2.) a demographic shift that has lowered the share of the school-aged population since the 1960s, following the baby boom. To allow for these we introduce changing college-educated shares for women, C_f, and men, C_m, and a variable capturing the ratio of school aged students to the working age population, S. In the model, the changing population shares effectively work as productivity shifters. In the basic model, these shifters enter production in the following way:[17]

$$\begin{aligned} z &= S^{-\gamma}(C_f N_T)^{\alpha} \\ N_M^m &= \mu^M S z = \mu^M S^{1-\gamma}(C_f N_T)^{\alpha} \\ Y_M &= (C_f N_f^M + (C_m + \frac{1}{wp^{C_m}}(1-C_m))S N_m^M) \end{aligned}$$

We assume a larger student population lowers human capital produced per student, with the elasticity γ. To take account of the selection effect associated with an increasing college education population, we also adjust the male market productivity by the college premium (wp^{C_m}).

Male teachers: Thirdly, we drop the stark assumption between males and females in the model. We allow college educate males to be allocated to teaching and for human capital

[17] The effect is slightly more involved once we additionally add male teachers.

to affect female market productivity. In particular, output of market production in the male and female college educated worker is now:

$$Y^M_{f,college} = A^M \mathcal{C}_f(\eta + Sz) \iint (1-\tau)(a^M)^{1-\hat{\tau}} \mathcal{M}_f(a^M, a^T) f(a^M, a^T) da^T da^M$$
$$Y^M_{m,college} = A^M \mathcal{C}_m(\eta + Sz) \iint (a^M) \mathcal{M}_m(a^M, a^T) f(a^M, a^T) da^T da^M$$

Where η capture market production by college educated workers unaffected by human capital.[18] We assume the underlying primative $F(a^T, a^M)$ is the same for

Complementarities in human capital production: Thirdly, we introduce non-perfect substitutability between teacher skill types. The teaching input is produced by the CES form production function:

$$N^T = A^{NT} \left(\left(\tilde{\mathcal{C}}_f \iint (a^T)^\rho \mathcal{T}_f(a^T, a^M) f(a^T, a^M) da^M da^T \right.\right.$$
$$\left.\left. + \tilde{\mathcal{C}}_m \iint (a^T)^\rho \mathcal{T}_m f(a^T, a^M) da^M da^T \right) \right)^{\frac{1}{\rho}}$$

where $\tilde{\mathcal{C}}_x = \mathcal{C}_x/(\mathcal{C}_f + \mathcal{C}_m)$, is the share of college educated females, $\tilde{\mathcal{C}}_f$, or males, $\tilde{\mathcal{C}}_m$. The motivation for this extension is the idea of within-school complementaries e.g. that the school production function cannot just be staffed by the lowest quality teachers. One example of this would be the importance of school principals or other senior teachers. We allow the data to inform the importance of these complementarities. In practice, such a specification is necessary to fit the level of the teacher wage premium in the data. This parameter is fixed over time.

Given these extension the constrained planner chooses a contract (w^T, a^T_{MIN}) to maximize:

$$\max_{a^T_{MIN}, w^T} \quad Y^M_f + Y^M_m + Y^H_f + Y^H_m$$

[18] This could also be thought of as the college premium and helps ensure an interior solution.

Subject to

$$
\begin{aligned}
Y_f^M =& A^M \mathcal{C}_f(\eta + Sz)\Bigg((1-\tau)\int_0^{a^T_{MIN}} \int_{\phi(\hat{a}_f^H)}^{\infty} (a^M)^{1-\hat{\tau}} f(a^T, a^M) da^M da^T \quad \text{(market prod. F)} \\
&+ (1-\tau)\int_{a^T_{MIN}}^{\infty} \int_{\phi(\hat{w}^T)}^{\infty} (a^M)^{1-\hat{\tau}} f(a^T, a^M) da^M da^T \Bigg) \\
&+ (1-\tau)(1-\mathcal{C}_f)(1-\mathcal{H}_f) A^M \mu^M \frac{1}{wp_f} Sz \\
Y_m^M =& A^M \mathcal{C}_m(\eta + Sz)\Bigg(\int_0^{a^T_{MIN}} \int_{\hat{a}_m^H}^{\infty} a^M f(a^T, a^M) da^M da^T \quad \text{(market prod. M)} \\
&+ \int_{a^T_{MIN}}^{\infty} \int_{\hat{w}^T}^{\infty} a^M f(a^T, a^M) da^M da^T \Bigg) \\
&+ (1-\mathcal{C}_m)(1-\mathcal{H}_m) A^M \mu^M \frac{1}{wp_m} Sz \\
Y_f^H =& \mathcal{C}_f a_f^H \int_0^{a^T_{MIN}} \int_0^{\phi(\hat{a}_f^H)} f(a^T, a^M) da^M da^T + a_f^H (1-\mathcal{C}_f)\mathcal{H}_f \quad \text{(home prod. F)} \\
Y_m^H =& \mathcal{C}_m a_m^H \int_0^{a^T_{MIN}} \int_0^{\hat{a}_f^H} f(a^T, a^M) da^M da^T + a_m^H (1-\mathcal{C}_m)\mathcal{H}_m \quad \text{(home prod. M)} \\
N^T =& A^{NT}\Bigg(\tilde{\mathcal{C}}_f \int_{a^T_{MIN}}^{\infty} \int_{-\infty}^{\phi(\hat{w}^T)} (a^T)^{\rho} f(a^T, a^M) da^M da^T \quad \text{(teacher input)} \\
&+ \tilde{\mathcal{C}}_m \int_{a^T_{MIN}}^{\infty} \int_{-\infty}^{\hat{w}^T} (a^T)^{\rho} f(a^T, a^M) da^M da^T \Bigg)^{\frac{1}{\rho}} \\
z =& A^T((C_f + C_m)N^T)^{\alpha} S^{-\gamma} \quad \text{(human capital)}
\end{aligned}
$$

The additional notation indicates the market barrier equivalent cutoff of female market skill in terms of the teacher wage: $\phi(\hat{w}^T) = (\hat{w}^T/(1-\tau))^{\frac{1}{1-\hat{\tau}}}$, the teacher wage in units of college educated human capital: $\hat{w}^T = w^T/(A^M[\eta + Sz])$ and home production in units of college educated human capital: $\hat{a}^H = a^H/(A^M[\eta + Sz])$.

It is important to note thatwe assume that the constrained planner is only able to choose a single teacher contract pair (w^T, a^T_{MIN}) and that this contract is not gender specific. This introduces an additional distortion due to a misallocation between males and females, when the female market barriers are positive. The unconstrained planners problem and first order conditions are presented in the Appendix.

6.1 Calibration

We now calibrate the model to take it to the data. We use a combination of external and internal calibration targets to discipline the results. Given a paucity of females in the data with less than a college education being observed as teachers, we calibrate the model to

targets in the data that capture allocations conditional on having some college education. The list of parameters is shown in Table 2.

Table 2: Calibrated parameters

Parameter		Value	Source
Externally calibrated			
Decreasing returns to teaching	α	0.880	Current expend.
Teaching and market correlation	$\rho_{M,T}$	0.903	O-Net
Standard deviation of teaching	σ_T	0.198	Teacher VA
Return to class size	γ	0.5	Rivkin et al (05)
Internally calibrated			
Mean teacher productivity	μ_T	-1.52	CPS
Standard deviation of market	σ_M	0.41	CPS
Teaching skill premia	ρ	1.96	CPS

The production of human capital $A_T S^{-\gamma}((C_f + C_m)N_T)^{\alpha}$ is Decreasing Returns to Scale. Therefore, we assume the output elasticity to the teaching input is equal to the teacher input factor share. We set α equal to the wage bill share of teachers in total school expenditures. More precisely, we use the National Center for Education Statistics (NCES) Public School Expenditure Tables and compute the share of current expenditures in total school expenditure.[19] We use the average between 1990 and 2020, which gives a value of $\alpha = 0.88$. The $S^{-\gamma}$ captures the decreasing returns to class size. These are typically estimated to be fairly low. We follow Rivkin et al. (2005) who find an upper bound for the class size effect of a 10-student change in class size generating benefits less than a one standard deviation increase in teacher quality. To implement this we set $\gamma = 10/25 \times \sigma_T$.

For $F(a_T, a_M)$, we assume the underlying productivity distributions are distributed jointly log-normally:

$$\begin{pmatrix} a^M \\ a^T \end{pmatrix} \sim \log\mathcal{N}\left(\begin{bmatrix} \mu^M \\ \mu^T \end{bmatrix}, \begin{bmatrix} \sigma_M^2 & \sigma_{M,T} \\ \sigma_{M,T} & \sigma_T^2 \end{bmatrix}\right) \tag{12}$$

where the correlation coefficient is $\rho_{T,M} = \sigma_{M,T}/(\sigma_M \sigma_T)$. We calibrate the correlation by using the task content of teaching versus non-teaching occupations. We use the O*NET-based tasked content of occupations of Acemoglu and Autor (2011). This defines the task content of an occupation under six broad composites.[20] We residualize log wages for workers with strong labor market attachment to get an estimate of worker productivity.[21]

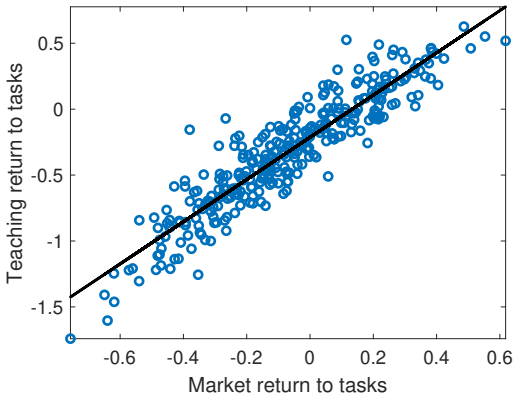
[19] Data is from Table 236.60. Total and current expenditures per pupil in fall enrollment in public elementary and secondary schools.

[20] These are non-routine cognitive analytic, non-routine interpersonal, routine cognitive, routine manual, non-routine manual physical, and non-routine manual interpersonal.

[21] Controls include a cubic polynomial in age and state, urban, race, and year fixed effects. The sample is restricted to workers earning more the $5,000, working more than 15 hours a week on average, and working more than 26 weeks last year.

Worker productivity is then regressed on the task content of occupations, separately for private sector occupations and those identified as teaching occupations. Productivity in both sectors is then predicted out of sample for all occupations, based on the occupations' associated task content. We then compute the predicted correlation in productivity between the private and teaching sectors across all occupations, weighted by worker shares. This gives a value for the correlation in productivities of $\rho_{M,T} = 0.903$. Conceptually, this estimate attempts to identify the correlation in the returns to an occupational bundle of tasks in the private and teaching sector.[22] A scatter plot of the predicted returns to tasks in the market and teaching sector is shown in Figure 12. The returns to underlying skills captured by task content across occupations inside and outside of teaching appear highly correlated.

Figure 12: Market and teacher productivity correlation



Notes:

The standard deviation of teacher productivity is set following estimates of teacher value added in the literature. Chetty et al. (2014) estimate an upper bound of 1.65 percent increase in earnings at age 28 from increasing teacher value added by one standard deviation for one grade.[23] Through the lens of the model we interpret the productivity of individual teachers to be distributed with standard deviation $\sigma_T = 12 \times 0.0165 = 0.198$ reflecting the improvement in labor market outcomes from increasing teaching productivity for the whole schooling period (12 grades). Finally, all aggregate productivity parameters (A_T, A_M) are set to 1.

The remaining parameters need to be internally calibrated within the model. We normalize average productivity in the market to $\mu_M = 1$. We still need to calibrate teacher productivity μ_T and the standard deviation of market productivity σ_M. This is a challenge because the female allocation may also still include a productivity barrier to work $\tau > 0$. Therefore, we calibrate a male version of the model under the assumption that $\tau = 0$ and that the primitives (μ_T, σ_M) in the absence of a barrier are the same for men

[22] As the correlation coefficient is unit-free, the actual scale of the distribution of either productivity is irrelevant.
[23] The lower bound with controls for parental characteristics and lagged incomes is estimated at 1.34 percent.

and women. More precisely, we calibrate (μ_T, σ_M) to match the standard deviation in residual male log wages which we estimate as 0.36 with a participation rate of 92.2 percent. We choose the average productivity of teachers, μ_T such that the share of male and female college workers allocated to teaching is equalized. Currently, only 8 percent of males in work are allocated to teaching, so we target a participation weighted $0.115 = (.80 \times .17 + .92 \times .8)/(2 \times .92)$. This is the allocation the social planner would choose when maximizing welfare, constrained by offering a teaching contract that features a flat wage and minimum skill requirement $\{w_T, a_{T,min}\}$. The low share of males working in teaching implies a low average estimate of the productivity of teaching of $\mu_T = -1.52$. If teaching on average was more productive the social planner would find it optimal to allocate a greater mass of workers to teaching. Finally, we choose ρ to target the current male teacher wage premium, defined as $w_T/(Y_M/F(N_M))$ in the model. In the data, this is -8.8 percent, which gives a parameter value of $\rho = 1.96$. This implies a significant benefit to allocating some high-productivity teachers to teaching.[24]

7 Declining barriers to work

7.1 Estimation and model fit

We can now take the model to the data. We use the structure of the model to estimate a sequence of parameters that capture the two productivity barriers facing females in the labor market and the value of home production$\{\tau_t, \hat{\tau}_t, a_{h,t}, a_{h_m,t}\}_{t=1970}^{2020}$. The model is a series of static equilibria. We divide the time period, into six ten-year periods plus an initial steady state which we assume is pre-1970. In each decade we target:

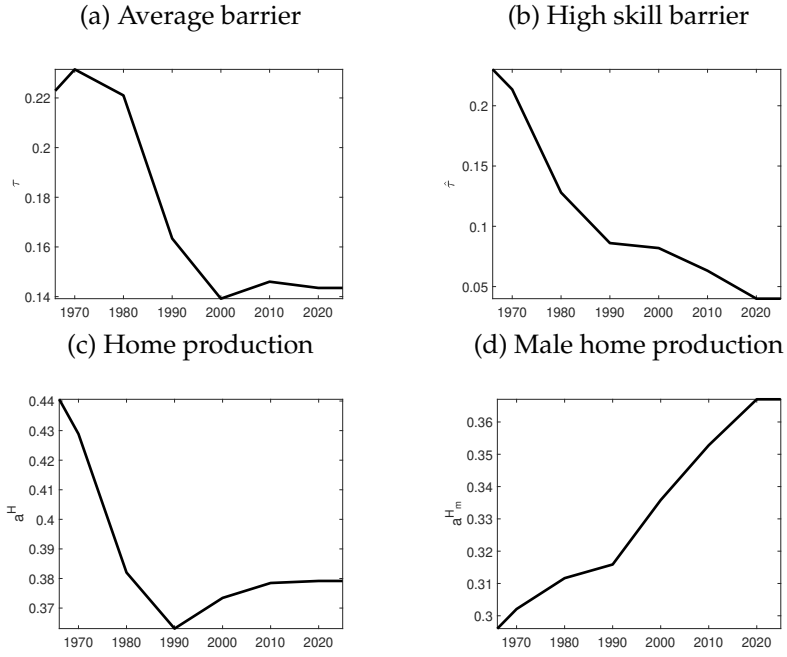
1. female labor force participation
2. the share of participating females working in teaching
3. the female teacher wage premium
4. male labor force participation

To uncover these parameters we assume each period the social planner implements the teaching contract $\{\{w_{T,t}, a_{T,min,t}\}\}_{t=1970}^{2020}$ to allocate females and males between home production, teaching, and the market. This is a constrained optimium, as the planner can only vary two instruments when setting this contract. Further, the same contract is available to female and male teachers and the planner cannot discriminate. For the annual targets, we use values in CPS for the sample of females with some college education (see Section 2).[25]

[24] An alternative way to think about this parameter is that it fattens the tail of the teaching skill distribution.
[25] In comparison to the Figure shown in Section 2, we residualize earnings in the CPS to try to ensure the targeted wage-premium does not include age effects, for example.

The estimates of these parameters are shown in Figure 13. The model suggests that barriers to work in the market sector have fallen substantially over time. In 1970, the estimated productivity barrier was 0.22 by 2000 it had fallen to 0.14, a fall of 38 percent. There were no further declines after 2000. This means in 2020, through the lens of the model women still faced a substantial barrier to market work. This is rationalized by the significantly larger share of college-educated women engaged in teaching as opposed to men. There was also a significant decline in the high-skill barrier. This declined from 0.23 in the 1970s to 0.04 today. This is rationalized by the declining teacher wage premium. Indeed, without the high-skill barrier it is challenging to match this feature of the data. Relative to the average barrier the high-skill barrier has continued to fall more recently. By contrast, the benefit of home production is much more stable. It fell slightly from 0.44 in 1970 to 0.38 in 1980. After that, it remained fairly stable. The model predicts that a large fraction of the increase in female labor force participation can be understood as the result of a decline in the productivity barriers facing females in the market sector, rather than an improvement or change in the technology of home production. For the men the model predicts a slight increase in the value of home production over this period, consistent with declining male participation.

Figure 13: Implied barriers to market production



Notes:

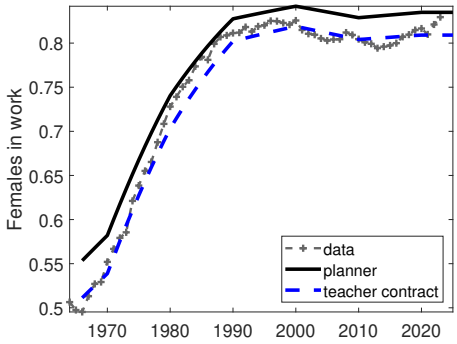
Figure 14 shows the allocation implied by the model for female and male participation, the share of teachers and the female teacher wage premium. The dashed blue line shows the allocation with the teacher contract. The line with crossed markers shows the data. For computational expediency, the model is calibrated to six 10-year average targets in the data and we then linearly interpolate the resulting barriers to market production and

solve for the optimal contract and implied allocation every year. Across all four panels, it is clear the model can replicate the observed allocations in the data. These were targets for uncovering the productivity barriers and value of home production, so the plots are essentially a test of the model fit. In particular, the model delivers both the increase in labor force participation and the decline in the share of teachers (conditional on participation).

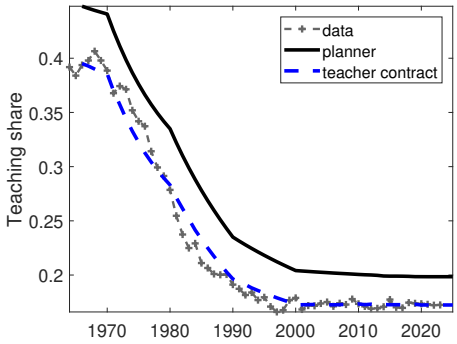
Figure 14 also shows what would happen under the optimal planner's allocation (planner). Here we see that under the optimally chosen teacher contract, the planner can implement an allocation very close to the unconstrained optimum. There are some differences. In the unconstrained equilibrium, there are slightly more females and males in work. The teaching share (the share of working females that are teachers) is also higher.

Figure 14: Female participation and allocation

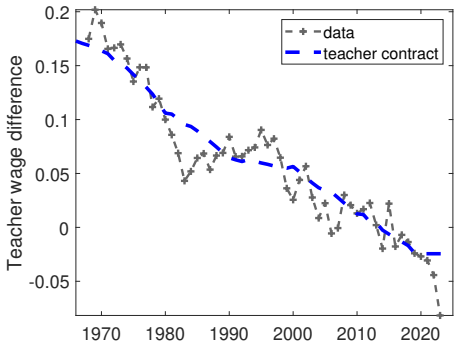
(a) Participation



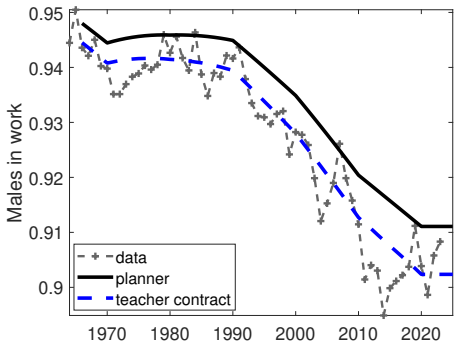
(b) Share of teachers (female)



(c) Wage premium



(d) Male Participation



Notes:

7.2 Optimal contract and aggregate predictions

7.2.1 Optimal contract

We now consider predictions of the model outside of the targeted moments and the implications for the aggregate economy. In addition to our teacher contract allocation, the model allows us to undertake counterfactuals under alternative contract assumptions. We compare the allocation under four alternative contracts:

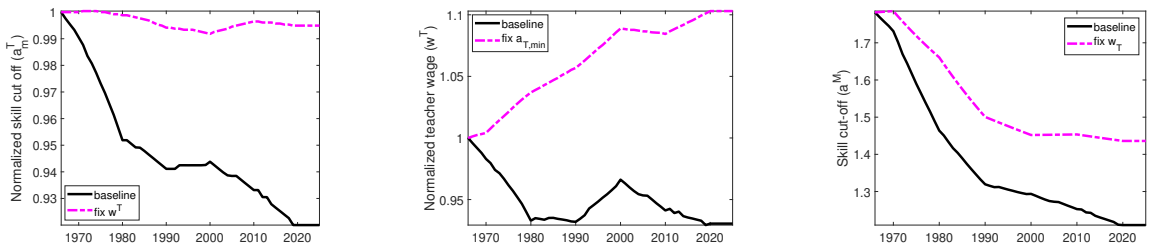
1. the planner's socially optimal allocation *(planner)*
2. the optimal teacher contract, fixing the wage at its initial level and choosing $a_{T,min}$ optimally *(fix w_T)*
3. the optimal teacher contract, fixing the skill requirement at its initial level and choosing w_T optimally *(fix $a_{T,min}$)*
4. fixing the teacher contract to the initial optimal contract *(fix both)*.

Figure 15 presents the optimal teacher contract between 1970 and 2020. The left panel shows the minimum skill cut off, $a_{T,min,t}$. Focusing on the optimal contract (black line), this falls with the decline in the barrier to market production. This ensures that while high-skilled female teachers are reallocating out of teaching into the marketplace, they are replaced by lower-skilled teachers. The middle panel shows the teacher wage, $w_{T,t}$. This also *also* initially declines, until the 1980s, before increasing again by about 25 percent above its 1980s level. While difficult to distinguish from business-cycle frequency dynamics, qualitatively this is not entirely different from the pattern we see in the aggregate data (Figure 3).

While the wage increases, this does not imply more productive females are moving into teaching as the productivity barrier is falling more quickly.[26] This can be seen in the right panel which shows the productivity cut-off implied by the teacher wage. It is monotonically falling. When the wage is fixed a smaller decline in the minimum skill level is required. Fewer high skilled teachers are being lost due to the lower wage, so fewer low quality teachers are needed. As can be seen from Panel b, the decline in the productivity cut-off is smaller. When the minimum skill cut-off is fixed, instead the teacher wage needs to rise to ensure those allocated to teaching does not fall, when the market sector becomes more attractive.

Figure 15: Optimal teacher contract

(a) Minimum skill (b) Wage (c) Productivity cut off



Notes:

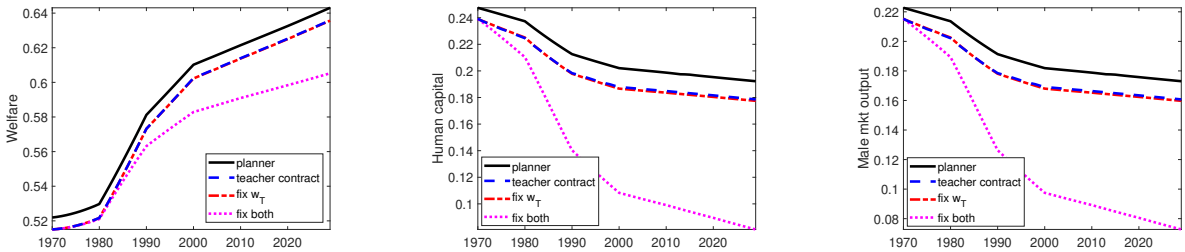
[26] Indifference, for these workers is characterized by $w_{T,t} = (1-\tau_t)a_M^{1-\hat{\tau}}$. So a fall in τ_t mechanically implies a decline in the cut off value of a_M and change in the mass of agents located at this indifference condition.

7.2.2 Aggregate predictions

Fixed college share: Before turning to our main results we briefly consider a model without male teachers and with fixed demographic shares. The implications for non-targeted aggregates are shown in Figure 16. The fall in labor market barriers substantially raises welfare in the economy (Panel a). The effective increase in the productivity of females in the market raises welfare by 22 percent. A second finding is that in this setting the planner can fairly closely implement this allocation with a restricted contract that allows either for (i) a change in both the minimum skill level and the wage paid or (ii) a change in the minimum skill level. These contracts generate a deviation from the first best allocation, of around 1 percent. This gap in welfare is fairly constant even as the barriers to work fall. Panels b shows the small welfare loss is due lower production of total human capital. Human capital itself falls by around 24 percent between 1970 and 2000. This then lowers male market output relative to the optimal allocation. Male output is about 6 percent lower than in the planner's allocations.

Figure 16: Predictions for aggregate economy (fixed college share)

(a) Welfare (b) Human capital (c) Male output



Notes:

Secondly, we observe distortions imposed by the fixed contract. Indeed, it is not true that aggregate outcomes in the economy are invariant to the contract offered to potential teachers. In particular, the contract that fixes the wage and minimum skill cut off of teachers at the initial values creates a large distortion. Under this contract, welfare is 8 percent lower in 2020. This contract reduces the number of teachers dramatically, as the barriers to working in the private sector decline. Average teacher quality falls by less, but the total teaching input falls substantially. The result is a large decline in aggregate human capital. This lowers male output and generates a substantial welfare loss.

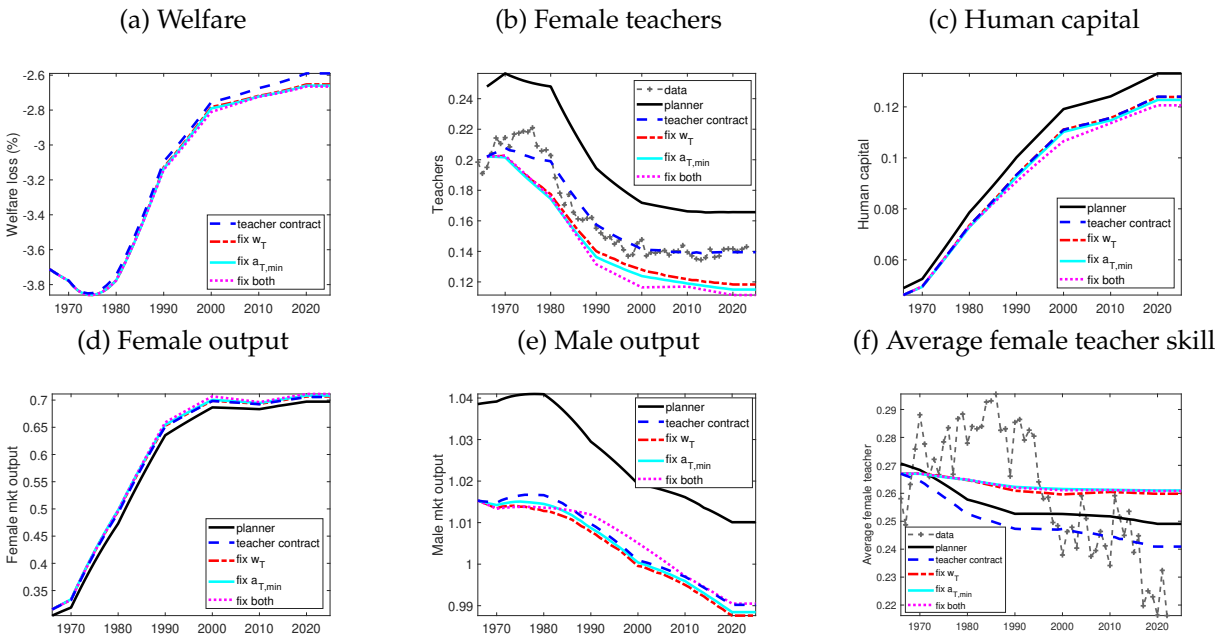
Baseline model: We now turn to the implications for non-targeted aggregates in our baseline economy. This features endogenous male teachers, an increasing college share of workers, and a fall in the student population.[27] Relative to the previous model, the changing college share (and to a lesser extent fall in student population) reduces the estimated average barrier at the beginning of the sample. The inital τ is 0.22, compared with 0.44 in the fixed college counterfactual. Two forces are at work. The increase in

[27]The time series for these trends are presented in the Appendix.

college educated women act like a productivity boost to the market sector which features linear production, relative to the teaching sector with decreasing returns. Thus part of the reallocation from teaching to market is explained by this force. The decline in the student population acts like a demand shifter. The 1970s when the student population was large was the time to allocate a larger share of the workforce to teaching. Thus some of the decline in market barriers is instead explained by this force.

As in the model with a fixed college share, the fall in labor market barriers substantially raises welfare in the economy. The effective increase in the productivity of females in the market and the increase in the college population raises welfare by more than 70 percent (Table 3, Column (1)). However, in contrast to the model above the welfare loss of the constrained contract is now much larger. While the constrained planner can fairly closely implement the allocation of the first best, the restriction imposed by the teacher contract does generate a sizeable welfare loss (Figure 17, Panel a). On average the welfare loss is worth around 3 percent. It reaches its maximum in the mid 1970. The magnitude of this welfare loss depends importantly on the misallocation that occurs between female and male teachers. If we instead allow only males to be teachers the average welfare loss is significantly smaller at around 1 percent (for example, Column (4)).

Figure 17: Predictions for aggregate economy



Notes:

Panel b shows that the supply of female teachers declines. This follows naturally from the fall in the targeted share. Further, we see that the unconstrained planner would choose a significantly higher fraction of the college educated female population to be teachers. Despite this decline in our baseline specification, human capital increases by around 150 percent between 1970 and 2000 (Panel c). This contrasts with the dynamics of the fixed

college share model where human capital falls. Here, the large increase in human capital is due to the greater supply of female college graduates (see Figure 11).

Consequently, as females are reallocated out of teaching and into the market and human capital increases there is a large increase in female market output (Panel d). This is the main source of welfare gains in the economy. From 1970 to 2020 female market output almost doubles. In contrast, male market output falls (Panel e). The decline in male output is due to both a fall in participation and a reallocation of some males into teaching. As human capital increases each unit of male market productivity produces more. Offsetting this, lower productivity men who were previously optimally allocated to the market sector are now reallocated to the teaching sector, substituting for females who move in the opposite direction. This leads to a decline in male output. Interestingly, the unconstrained planner would be able to increase male output by about 5 percent.

Despite the increase in human capital, average teacher quality declines over this time period by about 8 percent (Panel e). The fall is larger under the optimal contract than under the planner's first best allocation. It might be thought that this helps explain the decline in real wages and labor market outcomes of low-skilled men. The model clarifies that this pattern would occur in the first best allocation. Interestingly, when we compare the decline in average teacher skill in the model with the decline in the difference in years of education (between female teachers and the whole economy) the model does a good job of explaining the decline seen in the data. Overall, the model generates about 50 percent of the decline in this variable seen in the data.

Finally, we return to the distortions imposed by fixing a component of the teacher contract. Surprisingly in the baseline model, these tend to be fairly small. In Panel b, we see that under all fixed contract solutions, the share of female teachers is smaller. In the baseline model, the planner lowers the minimum skill requirement a^T_{MIN} or raises teacher wages w^T to offset the reallocation from teaching to the market sector that occurs due to higher market sector wages. In the absence of this, the share of females allocated to teaching falls further. This results in slightly lower human capital, particularly when both the minimum skill level and wage are fixed. Most notably, we see that the average female teacher's skill level falls by much less than under the optimal contract, as lower quality teachers are not permitted to replace those reallocating.

There are two reasons why fixed contracts are less distortionary in the baseline model relative to the model with fixed college. Firstly, female market productivity now also depends on human capital. Therefore, if human capital produced by the teaching sector falls, wages in the market sector decline. This provides an inbuilt stabilizer against teachers reallocating out of the teaching sector when the teacher contract is fixed. Secondly, there is an additional source of teacher input: male teachers. Due to the same mechanism, if

human capital declines males also face a greater incentive to reallocate to the teaching sector and substitute for female teachers.

Table 3: Summary of model predictions

		Baseline		Female teachers only			
				College share			Dynamic
		US	Sweden	fixed	exog.	endog.	
		(1)	(2)	(3)	(4)	(5)	(6)
Average wedge (τ)	(p.p.)	-8.8	-3.4	-23.9	-4.3	-38.5	-20.4
High skill wedge (τ_2)	(p.p.)	-17.4	-32.1	-9.6	-10.3	-40.1	-5.3
Welfare	(%)	73.3	107.5	21.2	207.5	1122.3	48.8
Teaching input (N_T)	(%)	-13.9	-39.4	-26.7	-26.0	-27.1	-17.1
Human capital	(%)	150.4	266.7	-23.9	125.6	120.6	167.8
Average female teacher	(%)	-7.5	-14.1	-8.1	-8.5	-9.9	-9.1
Minimum skill (a^T_{min})	(%)	-7.1	-12.2	-6.3	-6.6	-7.9	-15.2
Ave welfare loss	(%)	-3.2	-1.8	-1.3	-1.3	-8.4	-0.9
Max welfare loss	(%)	-3.9	-2.1	-1.5	-1.5	-40.2	-1.2

We summarize the findings of this section by comparing the change in the allocation between 1970 and 2020 across multiple specifications. In general, this shows the robustness of our key findings to various modeling choices. We include a baseline version calibrated to the Swedish economy, a version of the model where we allow the college share to be an endogenous variable (details in the Appendix), and a dynamic version discussed in the next section. The results are shown in Table 3. The decline in the average wedge varies across the specifications, as does the increase in welfare and human capital. However, the decline in the teaching input is fairly consistent, between 14 and 27 percent in the US calibrations. All three models predict a similar decline in average teacher quality of 8-10 percent. This is approximately 60 percent of the decline observed in the data. All the models also predict a falling minimum required teacher quality, implying the planner wants to optimally allow lower-skilled teachers into the profession to replace those reallocated to the private sector. The largest decline in the minimum teacher skill level is produced by the dynamic model with a forward looking planner. The models without a college choice imply a smallish 1-3 welfare loss due to the planner not being able to implement the optimal contract. As discussed above the introduction of male teachers raises the welfare loss of the teacher contract. The welfare loss is even larger in the model with endogenous college choice. This welfare loss occurs at the beginning of the sample when the implied barrier to undertaking a college education is estimated to be large.[28]

7.2.3 Predictions for the wage distribution

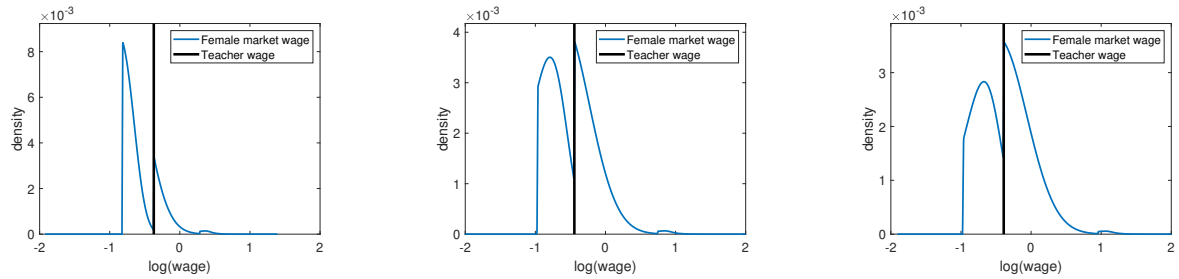
In Section 2.2, we documented a striking fact in the Swedish administrative data that when comparing the female teaching and private sector wage distributions we observe a shift in the mass of private sector wages from the left of the teacher wage distribution to the

[28]This follows mechanically from the significantly lower share of college-educated workers and females, in particular.

right. We now show that our model qualitatively generates the same pattern, In each year we generate a private sector wage distribution and compare it to the wage of the teacher contract. The results are shown in Figure 18. In our model there in no distribution of teacher wages as all teachers receive w^T. We see that in the first period (1970s) the majority of the wage mass in the private sector is to the left of the teacher wage. By 2020 there is a roughly even mass either side of the teacher wage. The model predicts that as the discriminatory barriers fall further the mass to the right will increase by more, capturing the stylized fact presented in the Swedish administrative data.

Figure 18: Changing wage density (female)

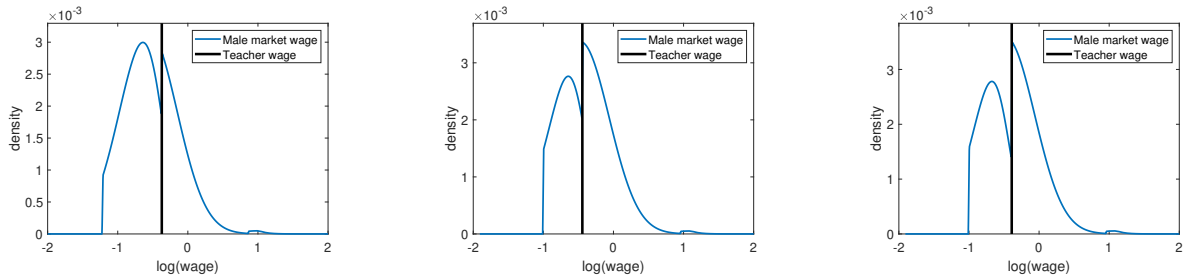
(a) 1970 (b) 2020 (c) $\tau = 0$



Similarly, the model also allows us to look at the effect on the male wage distributions. In the Swedish administrative data, these were fairly constant. Figure 19 shows that the baseline model with endogenous male teachers generates exactly the same pattern. While there is some shift in the distribution, the left mass of male wages moves nowhere near as much as the female wage density.

Figure 19: Constant wage density (male)

(a) 1970 (b) 2020 (c) $\tau = 0$



7.3 Forward looking dynamic model

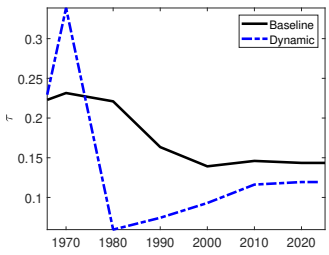
It might be unreasonable to assume, as we do in the baseline, that the change in demographic trends and falling market barriers were a persistent surprise to the planner. Further, occupational choices that require an early career time investment are not easily changed. Therefore, the existing allocation of workers may not easily be able to reallocate from one sector to another. To incorporate these frictions we solve a dynamic version of the model where the planner can only reallocate a fraction δ of the population each period. For this

exercise we restrict teachers to be female. Full details of the model are presented in the Appendix.

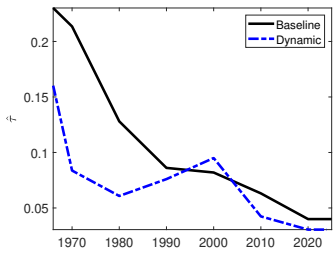
The dynamic model provides an important insight while leaving many other results unchanged (See Table 3). Figure 20 presents a comparison of the barriers in the dynamic model with the baseline. When it is difficult to reallocate females from the teaching sector to the market sector as market barriers decline. Then the static estimation will underestimate the pace and decline in the market barriers up to this point (Panel a). Observing lower barriers today and in the future the planner will choose a contract that lowers the share of females allocated to the teaching sector, but the share of teachers will depend on previously allocated women. This will tend to bias upward the estimate of current market barriers in a static setting. Further, the improvement in human capital is likely to be slower as it takes time for the benefit of the larger college sized female population to raise the teacher input (Panel c).

Figure 20: Dynamic model

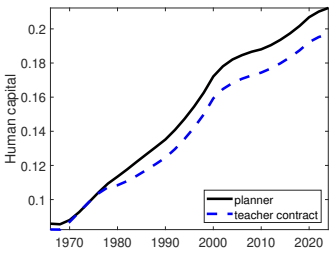
(a) Average barrier



(b) High skill barrier



(c) Human capital



7.4 Implications for future changes in teaching allocation

In our final exercise, we analyze how large the barriers are currently and what the implications are for the future under the circumstances of complete male-female parity. To do this we compute the economy setting both barriers equal to zero ($\tau = \hat{\tau} = 0$). We do this in our baseline specification with an exogenous increasing college share, but assume no further changes in the college share. The model implies there are still substantial welfare gains to accrue equal to a further 7 percent increase in welfare. Perhaps more interestingly, this implies a further 25 percent decline in the share of women in teaching and an optimal 2 percent decline in average teacher quality. Human capital would be fairly stable, with most of the decline in the female teacher share made up with increase demand for male teachers. In the economy with no discriminatory barriers, the teacher contract generates a 2.7 percent welfare loss relative to the first best. We also undertook the same analysis for the Swedish economy, calibrating the economy with the equivalent targets. The model predicts the economy is currently closer to an allocation with no wedges. Relative to the US, teaching is more dominated by females in Sweden and the teaching wage premium is also higher, but female participation is higher. These imply a smaller average barrier, but a

larger high-skill barrier. This means that average teacher quality has a further 10 percent to fall, more than double the change in the US, but the welfare gain is smaller.

	Welfare	Welfare distortion	Teacher share	Human capital	Average teacher
Unconstrained planner	6.7	0.0	-33.6	-0.3	-2.6
Teacher contract	6.9	-2.7	-26.2	-0.4	-2.0
Sweden contract	2.9	-1.7	-6.9	0.11	-10.1

8 Conclusion

This paper has documented and investigated trends in female participation and the reallocation in the occupation of high-skilled females from teaching to market production. It has shown that this trend is associated with a decline in the quality of teachers. In an estimated structural model, we find a significant decline in the barriers to market work faced by females. This has substantially raised welfare, while also lowering human capital production and male output. The model is able to reproduce the relative shift in the distribution of female private sector wages, that we observe in the Swedish administrative data. Contracting frictions in the teaching sector lower welfare by a fairly small 1 percent relative to the first best. However, the presence of rigid contracts imposes a larger welfare loss. In future work, we will extend the analysis to include a fully specified education choice, a dynamic effect of education on future human capital, and account for offsetting change in male teacher participation.

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4 of *Handbook of Labor Economics*, chapter 12, pages 1043–1171. Elsevier.

Alder, S. D., Duareva, Y., and Seshadri, A. (2023). The Allocation of Teaching Talent and Human Capital Accumulation. *Unpublished manuscript*.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–97.

Bacolod, M. P. (2007). Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality. *The Review of Economics and Statistics*, 89(4):737–751.

Chetty, R., Friedman, J. N., and Rockoff, J. (2016). Using Lagged Outcomes to Evaluate Bias in Value-Added Models. *American Economic Review*, 106(5):393–399.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9):2593–2632.

Goldin, C. (1990). *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Grönqvist, E. and Vlachos, J. (2016). One size fits all? The effects of teachers' cognitive and social abilities on student achievement. *Labour Economics*, 42(C):138–150.

Halldén, K. and Stenberg, A. (2023). Freedom of Career Choice? A Report on Gender Segregation in Education and Occupation. Technical report.

Hanushek, E. A. and Rivkin, S. G. (2012). The Distribution of Teacher Quality and Implications for Policy. *Annual Review of Economics*, 4(1):131–157.

Helen, McNally, S., Rolfe, H., Ruiz-Valenzuela, J., Savage, R., Vousden, J., and Wood, C. (2018). Teaching assistants, computers and classroom management: evidence from a randomised control trial. CEP Discussion Papers dp1562, Centre for Economic Performance, LSE.

Molinder, J. and Pihl, C. (2021). Women's work and wages in the sixteenth-century and Sweden's position in the "Little divergence". Lund Papers in Economic History 227, Lund University, Department of Economic History.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.

A Endogenous college model

- An individual choose college before knowing ability in both sectors
- Value of college:

$$V_C = Y_M^w + \tilde{N}_T w^T + Y_H - \mathcal{C} + \xi$$

  - $\mathcal{C}$: cost of college
  - ξ: Type I extreme value distributed with standard deviation σ_C
- Value of non-college, V_{NC}, time-varying reflecting non-college female wage penalty
- Share of females with college education:

$$C = \frac{\exp(V_C/\sigma_C)}{\exp(V_C/\sigma_C) + \exp(V_{NC}/\sigma_C)}$$

B Dynamic model

B.1 Assumptions

- Each period δ of households can be reassigned to T, M, or H. Call this a cohort i.
- Planner and households discount future with discount rate β.
- There is a path of $\{\tau_t, \hat{\tau}_t, a_t^H\}_{t=0}^{\infty}$ known to planner and household.
- The planner takes the beginning period stock: $\{N_{T,t-1}, Y_{MF,t-1}, Y_{H,t-1}\}$ as given.
- For the teacher contact: a teacher assigned in period i they must be above the time/cohort specific cut-off $a_{MIN,i}^T$.
- Until reassigned a teacher assigned in period i receives w_i each period, determined in the period they were assigned (i.e. teacher wages are cohort specific).
- In the market, a female assigned in period i faces the prevailing average distortion τ_t, *but the high-skill distortion* in the period they were assigned $\hat{\tau}_i$.
- A female assigned to home production in period i produces with a_i^H until reassigned.

In period t the market production by females is:

$$\begin{aligned}
Y_{MF,t} =&(1-\tau_t)A_{M,t}\sum_{i=-\infty}^{t}\mathcal{C}_i\iint(a^M)^{1-\hat{\tau}_i}\mathcal{M}^i(a^T,a^M)\delta F(da^T,da^M)(1-\delta)^{t-i}\\
=&(1-\tau_t)A_{M,t}\mathcal{C}_t\iint(a^M)^{1-\hat{\tau}_t}\mathcal{M}^t(a^T,a^M)\delta F(da^T,da^M)+\\
&(1-\tau_t)A_{M,t}\sum_{i=-\infty}^{t-1}\mathcal{C}_i\iint(a^M)^{1-\hat{\tau}_i}\mathcal{M}^i(a^T,a^M)\delta F(da^T,da^M)(1-\delta)^{t-i}\\
=&(1-\tau_t)A_{M,t}\mathcal{C}_t\iint(a^M)^{1-\hat{\tau}_t}\mathcal{M}^t(a^T,a^M)\delta F(da^T,da^M)+(1-\delta)\frac{A_{M,t}(1-\tau_t)}{A_{M,t-1}(1-\tau_{t-1})}Y_{MF,t-1}
\end{aligned}$$

Now we make the teacher input:

$$\begin{aligned}
N_{T,t} =&A_{NT}\Big(\sum_{i=-\infty}^{t}\mathcal{C}_i\iint(a^T)^{\rho}\mathcal{T}^i(a^T,a^M)\delta F(da^T,da^M)(1-\delta)^{t-i}\Big)^{\frac{1}{\rho}}\\
N_{T,t}^{\rho} =&A_{NT}^{\rho}\mathcal{C}_t\iint(a^T)^{\rho}\mathcal{T}^t(a^T,a^M)\delta F(da^T,da^M)+\\
&A_{NT}^{\rho}\sum_{i=-\infty}^{t-1}\mathcal{C}_i\iint(a^T)^{\rho}\mathcal{T}^i(a^T,a^M)\delta F(da^T,da^M)(1-\delta)^{t-i}\\
=&A_{NT}^{\rho}\mathcal{C}_t\iint a^T\mathcal{T}^t(a^T,a^M)\delta F(da^T,da^M)+(1-\delta)N_{T,t-1}^{\rho}\\
N_{T,t} =&\Big(A_{NT}^{\rho}\mathcal{C}_t\iint a^T\mathcal{T}^t(a^T,a^M)\delta F(da^T,da^M)+(1-\delta)N_{T,t-1}^{\rho}\Big)^{\frac{1}{\rho}}
\end{aligned}$$

And teacher production

$$Y_{MM,t} =\mu_t^Z A_{M,t}A_{T,t}(N_{T,t})^{\alpha}$$

The envelope condition for teacher input is now:

$$\begin{aligned}
\beta V_{Y_{N_T}} =&\beta(1-\delta)\alpha A_{M,t+1}A_{T,t+1}\mu_{t+1}^Z\mathcal{C}_{t+1}^{\alpha}(N_{T,t+1})^{\alpha-1}\Big(\frac{N_{T,t+1}}{N_T}\Big)^{1-\rho}+\\
&\beta^2(1-\delta)V_{Y_{N_{T,t+1}}}\Big(\frac{N_{T,t+1}}{N_T}\Big)^{1-\rho}\\
=&\sum_{i=1}^{\infty}[\beta(1-\delta)]^i\Big(\frac{N_{T,t+i}}{N_T}\Big)^{1-\rho}\alpha A_{M,t+i}A_{T,t+i}\mu_{t+i}^Z\mathcal{C}_{t+i}^{\alpha}(N_{T,t+i})^{\alpha-1}
\end{aligned}$$

The first order conditions for the constrained planner are:

$$A_M(1-\tau)\mathcal{C}\int_{\phi(a^H)}^{\phi(w^T)}(a^M)^{1-\hat{\tau}}\delta F(a_{MIN},da^M)\left(1+\sum_{i=1}^{\infty}[\beta(1-\delta)]^i\frac{A_{M,t+i}(1-\tau_{t+i})}{A_{M,t-1+i}(1-\tau_{t-1+i})}\right)$$
$$-\left(\sum_{i=0}^{\infty}[\beta(1-\delta)]^i\alpha\frac{A_{NT}^{\rho}}{\rho}A_{M,t+i}A_{T,t+i}\mu_{t+i}^{Z}\mathcal{C}_{t+i}^{\alpha}(N_{T,t+i})^{\alpha-\rho}\right)\int_0^{\phi(w^T)}(a_{MIN})^{\rho}\delta F(a_{MIN},da^M)+$$
$$\frac{a^H\mathcal{C}}{1-\beta(1-\delta)}\int_0^{\phi(a^H)}\delta F(a_{MIN},da^M)=0$$

and:

$$-A_M(1-\tau)\mathcal{C}\frac{w^T}{\boldsymbol{\Sigma}}\int_{a_{MIN}}^{\infty}\delta F(da^T,\phi(w^T))\left(1+\sum_{i=1}^{\infty}[\beta(1-\delta)]^i\frac{A_{M,t+i}(1-\tau_{t+i})}{A_{M,t-1+i}(1-\tau_{t-1+i})}\right)+$$
$$\left(\sum_{i=0}^{\infty}[\beta(1-\delta)]^i\alpha\frac{A_{NT}^{\rho}}{\rho}A_{M,t+i}A_{T,t+i}\mu_{t+i}^{Z}\mathcal{C}_{t+i}^{\alpha}(N_{T,t+i})^{\alpha-\rho}\right)\int_{a_{MIN}}^{\infty}(a^T)^{\rho}\delta F(da^T,a_{MIN})=0$$

The indifference condition for the unconstrained planner are:

$$\left(\sum_{i=0}^{\infty}[\beta(1-\delta)]^i\alpha\frac{A_{NT}^{\rho}}{\rho}A_{M,t+i}A_{T,t+i}\mu_{t+i}^{Z}\mathcal{C}_{t+i}^{\alpha}(N_{T,t+i})^{\alpha-\rho}\right)(a^T)^{\rho}=$$
$$A_M(1-\tau)\mathcal{C}(a^M)^{1-\hat{\tau}}\left(1+\sum_{i=1}^{\infty}[\beta(1-\delta)]^i\frac{A_{M,t+i}(1-\tau_{t+i})}{A_{M,t-1+i}(1-\tau_{t-1+i})}\right)$$
$$\left(\sum_{i=0}^{\infty}[\beta(1-\delta)]^i\alpha\frac{A_{NT}^{\rho}}{\rho}A_{M,t+i}A_{T,t+i}\mu_{t+i}^{Z}\mathcal{C}_{t+i}^{\alpha}(N_{T,t+i})^{\alpha-\rho}\right)(a^T)^{\rho}=\frac{a^H}{1-\beta(1-\delta)}$$
$$A_M(1-\tau)\mathcal{C}(a^M)^{1-\hat{\tau}}\left(1+\sum_{i=1}^{\infty}[\beta(1-\delta)]^i\frac{A_{M,t+i}(1-\tau_{t+i})}{A_{M,t-1+i}(1-\tau_{t-1+i})}\right)=\frac{a^H}{1-\beta(1-\delta)}$$

Given a path of $\{\mathcal{C}_t,\mu_t^Z,N_{T,t},\tau_t,\hat{\tau}_t,a_t^H\}_{t=1}^T$ and an initial $N_{T,-1}$ a solution satisfies the first order conditions and implies the expected path $\{N_{T,t}\}_{t=1}^T$ satisfies market clearing $N_{T,t}=N_{T,t}^*(\mathcal{T}^t,\mathcal{M}^t)\ \forall t$.

B.2 Solution method

From the intial 1970's steady state guess a path $\{\tau_t,\hat{\tau}_t,a_t^H\}_{t=1}^T$ and update until we match time series of working share, teacher share and teacher wage premium.