Education Expansion, Technological Change, and the Decreasing Wage Premium

Pauline Corblet *

October 13, 2021

JOB MARKET PAPER

Click here for the latest version

Abstract

The wage gap between educated and uneducated workers has grown quickly in the 1970s and the 1980s in the US, despite the uninterrupted rise in the relative supply of educated workers. Since the 2000s however, the education wage premium has stagnated in many western economies. Using an exhaustive matched employer-employee dataset, I study the case of the Portuguese labor market, where the high school wage premium has stalled since the late 1980s and even decreased in more recent years. This decrease has occurred in the context of rapid education expansion and the economy's transition from manufacturing to services. I examine the heterogeneity of the graduate wage premium evolution to age, as its decrease is particularly sharp for young workers. I investigate three channels to explain this decrease: the rise of the educated workers relative supply, changes in demand brought by technological change, and sorting between workers and firms. To do so, I build a model of one-to-many matching, which uses the most recent advances in matching theory and computational matching to make predictions both on wages and matching. The model affords flexibility in identification that lets me estimate sorting between workers and firms as well as varying elasticities of substitution over time. I find evidence of skill-biased technological change that has increased the demand for educated workers. However the magnitude of technological change varies by industry, and I find that young educated workers favor industries where technological change is most modest.

*Sciences Po, pauline.corblet@sciencespo.fr

I thank my advisor Alfred Galichon for his valuable guidance and support. I am indebted to Francesco Franco who generously helped me obtain the data used here. I also thank Jean-Marc Robin, Florian Oswald, Kerstin Holzheu, Zsofia Barany as well as the participants to Sciences Po PhD and Equiprice seminars for their comments and feedback. All remaining errors are mine.

1 Introduction

It is a well-established fact that high school and university graduates experience higher wages than their less-educated peers in most labor markets. The education wage premium is especially documented in the US, where it has increased quickly in the 1970s and 1980s, resulting in the widening of the wage structure and rising inequality between college graduates and their less-educated peers (Goldin and Katz (2008)). Over the same period, the relative supply of college graduates over high school graduates has increased by 4.5% per year between 1967 and 1982 (Card and DiNardo (2002)). The prevailing reason given to the coexistence of these two facts is the advent of skill-biased technological change (SBTC), whereby the rise of new technologies, and in particular the use of computers, has increased firm demand for graduate workers, pressuring their wage upwards (Katz and Murphy (1992)). Because computers complement workers performing non-routine, cognitive tasks, firms' growing reliance on computers increases demand for educated, high-skilled workers (Autor et al. (2003)). However, if the SBTC hypothesis has proven a powerful explanation for the quick increase in graduate wage premium of the 1970s and 1980s, it is less clear if it can rationalize the subsequent slow down of both graduate wage premium and graduate supply in the 1990s, when the use of computers became prevalent (Card and DiNardo (2002), Beaudry and Green (2005)).

This paper studies the case of Portugal between 1987 and 2017. Portugal is a particularly relevant example of rapid supply and demand changes on the labor market: it entered the European Union in 1986, which fuelled its economy's transition from being dominated by manufacturing (50% of the labor force employed in 1987), to services (30% of the labor force employed in 2017). Meanwhile, only 10% of its employed labor force held a high school degree in 1987, a percentage that has risen to 50% in 2017. As a point of comparison, the percentage of high school graduates in the US workforce has gone from 75% to 90% over the same period[1]. The proportional increase of high school graduates in Portugal is more extensive and starts from a much lower presence of high school graduates on the labor market than in the US. In this respect, it is closer to the change in university graduates on the US labor market (from 20% to 35% over the same period). Graduating from high school has become much more common in Portugal over the last thirty years, but it is only in 2007 that high school graduates started representing the majority of young workers between 25 and 30. In 2017, 32% of the young workers between 25 and 30 still do not hold a high school degree. Meanwhile, university graduates in Portugal represented less than 3% of the employed labor

[1] Percentages computed over workers aged more than 25, Census data

force in 1987, and about 19% in 2017. Because the share of university graduates remain small for most of the period (it only reaches 10% in 2005), and because graduating from high school is still quite uncommon over most of the period I study, I consider a high school degree to be a differentiating signal in skill on the Portuguese labor market, much as a college degree is on the US labor market, and I choose to focus on the high school wage premium in this paper.

Since the 2000s, the high school wage premium decreases in Portugal, and the college wage premium stagnates in the US. Various explanations for this stagnation have been put forward: Beaudry et al. (2015) argue that the demand for cognitive skills has decreased since the early 2000s, pushing graduate workers down the job ladder. Valletta (2016) also emphasizes the role of job market polarization, i.e. the shift away from middle-skilled occupations, on college graduates' wages (as opposed to postgraduates). On the contrary, Blair and Deming (2020) examine job vacancy data and find that demand for skills has increased since the Great Recession. They explain the stagnating graduate wage premium by an increase in the supply of new graduates after 2008. They are backed by Hershbein and Kahn (2018) who show that the Great Recession has accelerated skill-biased technological change. Beyond diverging conclusions, this literature relies on the concept of a race between education and technology (RBET), a term first coined by Tinbergen (1974): both relative supply of graduates and demand for more educated workers have grown since the early 1970s, but their different rates of growth have caused changes in the wage structure. In Portugal, I find evidence of sustained, if modest, skill-biased technological change between 1987 and 2017, which indicates that the decreasing high school wage premium is rather due to the relative increase in the supply of high school graduates.

The graduate wage premium stagnation is observed together with rising inequality within education groups (Autor et al. (2020)). In particular, the overall high school wage premium in Portugal decomposes into very different wage premia by age groups. Younger high school graduates experience a lower wage premium compared to their uneducated peers than senior high school graduates do. This is true from 1987 to 2017, and the differential between young and senior high school wage premium widens over time. There can be three reasons behind the age heterogeneity in wage premia: the first is that the increase of relative high school graduates is mechanically faster among young workers than senior workers in an economy that knows an accelerating education expansion. The second is that technological change could be biased towards senior workers. The third has to do with how workers and firms sort on the labor market: if young workers favor firms in which technological change is not or only weakly biased towards high school graduates, their wage premium mechanically decreases compared to senior workers. This paper explores the supply and demand mechanisms leading to the general decrease in high school wage premium in Portugal between 1987 and 2017. It also investigates the age heterogeneity in this decrease by considering all three channels aforementioned. I find evidence of modest high school biased technological change, but no evidence of age biased technological change. I also find that young high school graduates favor industries in which technological change is the smallest. To my knowledge, this paper is the first to explore this channel.

To evaluate the weight of education's and technology's impact on the decreasing Portuguese high school wage premium, I extend the classical supply and demand framework developed by Katz and Murphy (1992) and Card and Lemieux (2001) by building a model of two-sided one-to-many matching. Using a model has two main benefits: first, it lets me investigate sorting between workers and firms, and I can consistently estimate both worker preferences towards firm industry and varying measures of technological change by industry. Second, I am able to use both wage and matching predictions to fit the data. This affords more flexibility in estimation, and I am able in particular to identify varying elasticities of substitution between worker education levels and age groups. In the model, several workers match with a single firm, and the surplus created by the match depends on the firm type workers' type, through the hired workforce composition. Utility is transferable and firms pay market-clearing wages to their workforce. Firms seek to maximize profit, and workers to maximize their utility, which is additive in wage and amenity that accounts for worker preferences towards some types of firms. Both firms and workers experience an additional random shock that accounts for unobservables in their type. The model is close to the seminal work of Kelso and Crawford (1982) on one-to-many matching. However, Kelso and Crawford (1982) make the crucial assumption that workers are gross substitutes (i.e. that there are no complementarities in firm preferences), an assumption I must do without to obtain a model that I can bring to the data. The matching literature has long contended with the gross substitute assumption (Hatfield and Milgrom (2005)), until a recent work by Che et al. (2019) in the non-transferable utility case and Azevedo and Hatfield (2018) in the transferable utility case, among others. I build on their work and rely on a continuum of workers assumption to show there exists an equilibrium in the model I develop.

In order to get meaningful predictions from my model, I use the framework developed by the hedonic literature (Ekeland et al. (2004)), and in particular the tools originating from the optimal transport literature and brought to matching by Dupuy and Galichon (2014), Galichon (2016) and Galichon and Salanie (2020), which I generalize from a one-to-one to a one-to-many matching set up.

Using an administrative matched employer-employee dataset provided by the Portuguese National Statistics Institute, I weight the effect of education and technology on the high school wage premium by assuming a nested Constant Elasticity of Substitution production function. The production function is nested over two levels, age and education, and takes labor inputs by education level and age groups. The analysis in two steps: I first run reduced-form estimations on the classic model developed by Katz and Murphy (1992) and Card and Lemieux (2001), where workers have no preferences, labor input is assumed to be exogenous and wages to be competitive. I find that high school graduates and nongraduates are highly substitutable, and the economy experiences technological change over the period. In a second step, I estimate the one-to-many matching model, using Generalized Least Squares on observed wages and maximum likelihood on observed matching. I obtain estimates for the production function that all vary in time, including the elasticities of substitution, as well as estimates for worker preferences. I find evidence of heterogeneous skill-biased technological change magnitudes depending on industry, as well as increasing substitutability between high school graduates and nongraduates. I also find that younger high school graduates display higher preferences towards industries that go through slower technological change, such as services.

Section 2 describes the evolution of the Portuguese high school wage premium between 1987 and 2017. Section 3 lays out the production framework. Section 4 estimate this framework with the simple model of Card and Lemieux (2001), and section 5 describes the one-to-many matching model. Section 6 discussed the model's identification and estimation, and section 7 presents estimation results.

2 The High School Wage Premium

2.1 Data Description

The *Quadros de Pessoal* dataset offers an exhaustive snapshot of the Portuguese labor market every year from 1987 to 2017. It covers all employees in the private sector (except domestic workers), and provides information on their age and highest degree obtained, as well as their monthly wage and hours worked. To compute the high school wage premium by age, I part the worker population into two groups: those who did not graduate from high school, and those who did. I also categorize workers into three age groups: young workers (from 16 to

35 years old), middle aged workers (from 36 to 50 years old), and senior workers (from 51 to 68 years old). I only consider full time employees, that is, workers that are neither part time workers (approximately 10% of the observations) nor self-employed, in unpaid family care, or in other forms of unemployment (less than 1% of the observations). I compute real hourly wage as the ratio of monthly wage over monthly hours, controlling for inflation and clean out the lowest 1% and highest 99% hourly wage percentiles.

2.2 Wage Premia by Age

$$\log w_{it} = \sum_{a_i \in \{y,m,s\}} \mathbb{1}_{[\text{HS graduate}_i]} \beta_{at} + g_i + r_{it} + d_{it} + u_{it} \tag{WP}$$

Where each individual i at time t earns wage w_{it}. $\mathbb{1}_{[\text{HS graduate}_i]}$ is an indicator function equal to 1 if i graduated from high school, and 0 otherwise. a_i is individual i's age belonging to either of the three categories y, m or s. g_i, r_{it} and d_{it} are gender, region and industry fixed effects. β_{at} is the yearly graduate wage premium, differentiated by age: it measures how much more (in %) a high school graduate earns than a non high school graduate, over time and by age bin.

Figure 1 shows the change in estimated high school wage premium over time for each age bin, along with 5% confidence intervals.

Figure 1: Estimated high school graduate wage premium over non graduates of same age

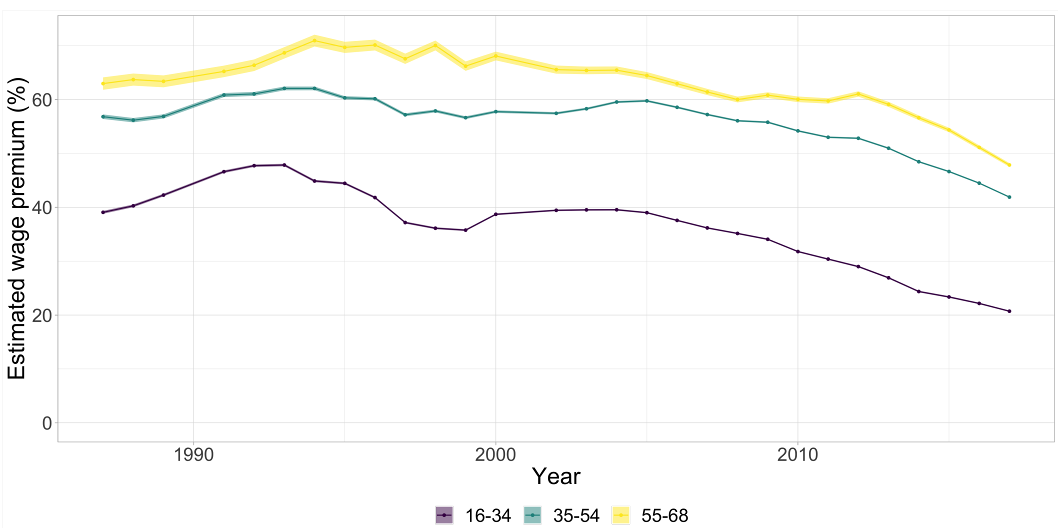
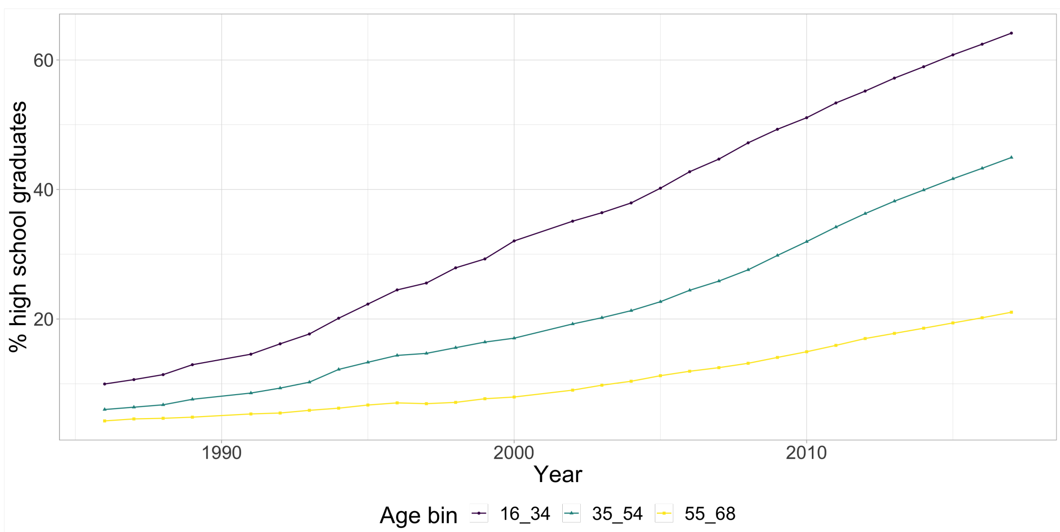


Figure 1 shows that high school wage premia differs by age bin: it is much higher (between 60% and 80%) for senior workers than for younger workers (between 40% and 20%). It also shows that it has decreased for all age bins between 1987 and 2017. But the extent of this decrease is different depending on age: senior workers lose only about 17p.p in high school wage premium over the period, while young workers lose slightly less than 50p.p and middle ages workers lose almost 30p.p. The decrease also starts earlier for young and middle aged workers (around 1995) than it does for senior workers (around 2000).

Relative supply of high school graduates versus non graduates has increased almost linearly for all age categories between 1987 to 2017, as evidenced by figure 2. Because high school enrolment grows every year, young workers are most impacted by this growth, and gain more than 50p.p of high school graduates over the period, while senior workers gain only about 20p.p.

Figure 2: Percentage of high school graduates in Portuguese labor force by age



3 Production Framework

Assume each firm on the labor market produces according to a nested Constant Elasticity of Substitution (CES) production function. Each firm belongs to some type $y \in \mathcal{Y}$, which determines the parameters of its production function. Firm production is a function of labor input by education level and age bin in any given year t:

$$\gamma_y(t) = \left[(\theta_H^y(t) H(t))^{\frac{\sigma(t)-1}{\sigma(t)}} + (\theta_L^y(t) L(t))^{\frac{\sigma(t)-1}{\sigma(t)}} \right]^{\frac{\sigma(t)}{\sigma(t)-1}} \tag{1}$$

Production γ's outer nest involves three parameters that may all vary with time: $\sigma(t)$, $\theta_H^y(t)$, $\theta_L^y(t)$ and two aggregate inputs $H(t)$ for high school graduates and $L(t)$ for non high school graduates. $\sigma(t) \in (0, \infty)$ is the elasticity of substitution between education levels, it is greater than one if graduates and non graduates are gross substitutes, and smaller than one if they are gross complements. $\sigma(t)$ is assumed to be the same across firm types. $\theta^y_H(t)$, $\theta^y_L(t) \in [0, \infty)$ are graduates and non graduate's productivity parameters. Both parameters may vary by firm type y. Aggregate labor inputs $H(t)$ and $L(t)$ are defined as

$$\begin{aligned} H(t) &= \left[\sum_{a \in \{y,m,s\}} \lambda^y_{a,H}(t) H_a(t)^{\frac{\tau^H(t)-1}{\tau^H(t)}} \right]^{\frac{\tau^H(t)}{\tau^H(t)-1}} \\ L(t) &= \left[\sum_{a \in \{y,m,s\}} \lambda^y_{a,L}(t) L_a(t)^{\frac{\tau^L(t)-1}{\tau^L(t)}} \right]^{\frac{\tau^L(t)}{\tau^L(t)-1}} \end{aligned} \tag{2}$$

Where $H_a(t)$ and $L_a(t))$ are labor inputs provided to the firm by high school graduates and non graduates in age category a. Labor input is measured in number of full-time equivalent workers. $\lambda^y_{a,H}(t)$, $\lambda^y_{a,L}(t) \in [0, \infty]$ are age productivity parameters, differentiated by education level and firm type. τ_H, $\tau_L \in (0, \infty)$ are elasticities of substitution between age categories. They vary by education level but are the same across firm types.

The production function is closed to the one used by Katz and Murphy (1992), and Card and Lemieux (2001): it assumes imperfect substitution and varying productivity in the tasks performed by different education levels and age categories. Capital is not included as an input, but may impact productivity parameters through firm type: if two firm types use different levels of capital in relation to education levels, it is reflected in the levels of $\theta^y_H(t)$ and $\theta^y_L(t)$. Unbiased technological change that increases all workers productivity results in an increase in both $\theta^y_H(t)$ and $\theta^y_L(t)$. Technological change may be biased towards an education level if its productivity increases faster than the other's. The present production function also allows more flexibility than Card and Lemieux (2001) by letting elasticities of substitution and age productivity vary in time.

Production assumes constant returns to scale. Note that it is homogeneous of degree one, and therefore two functions parametrized with θ and λ or $c \times \theta$ and $\frac{\lambda}{c}$ are equivalent. To distinguish between these versions, I impose normalization condition:

$$\sum_a \lambda^y_{a,H} = \sum_a \lambda^y_{a,L} = 1 \quad \forall y \tag{3}$$

4 Reduced Forms

Katz and Murphy (1992) and Card and Lemieux (2001) have shown that the CES production function parameters are identified from assuming that labor is optimally supplied to the economy and that wages are competitive, that is assuming that in each year t a representative firm solves

$$\max_{H_a, L_a} \gamma(t) - \sum_{a \in \{y,m,s\}} H_a w_{H,a} - \sum_{a \in \{y,m,s\}} L_a w_{L,a} \tag{4}$$

Where $\gamma(t)$ is the CES production function described in section 3 with no dependence on firm type, as in this set up I assume a single representative firm. I also assume in this section that elasticities of substitution τ^H, τ^L, σ, as well as age productivity parameters $(\lambda_{H,a})_a$, $(\lambda_{L,a})_a$ do not vary with time. Wages are competitive and equal to marginal productivity:

$$\begin{aligned} w_{H,a}(t) &= \lambda_{H,a}^{\frac{\tau_H - 1}{\tau_H}} H_a(t)^{-\frac{1}{\tau^H}} \times \theta_H(t)^{\frac{\sigma-1}{\sigma}} H(t)^{\frac{1}{\tau_H} - \frac{1}{\sigma}} \times \gamma(t)^{\frac{1}{\sigma}} \quad \forall a \in \{y, m, s\} \\ w_{L,a}(t) &= \lambda_{L,a}^{\frac{\tau_L - 1}{\tau_L}} L_a(t)^{-\frac{1}{\tau^L}} \times \theta_L(t)^{\frac{\sigma-1}{\sigma}} L(t)^{\frac{1}{\tau_L} - \frac{1}{\sigma}} \times \gamma(t)^{\frac{1}{\sigma}} \quad \forall a \in \{y, m, s\} \end{aligned} \tag{5}$$

Which results in relative wage equations:

$$\begin{aligned} \log\left(\frac{w_{H,a}(t)}{w_{H,a'}(t)}\right) &= \frac{\tau_H - 1}{\tau_H} \log\left(\frac{\lambda_{H,a}}{\lambda_{H,a'}}\right) - \frac{1}{\tau^H} \log\left(\frac{H_a(t)}{H_{a'}(t)}\right) \\ \log\left(\frac{w_{L,a}(t)}{w_{L,a'}(t)}\right) &= \frac{\tau_L - 1}{\tau_L} \log\left(\frac{\lambda_{L,a}}{\lambda_{L,a'}}\right) - \frac{1}{\tau^L} \log\left(\frac{L_a(t)}{L_{a'}(t)}\right) \end{aligned} \tag{6}$$

Restricting $(\lambda_{H,a})_a$, $(\lambda_{L,a})_a$'s variation in time, and adding a stochastic shock to account for measurement errors in observed wage and hours worked, relative age productivity and age elasticities of substitution can therefore be estimated by ordinary least squares through equations:

$$\begin{aligned} \log\left(\frac{w_{H,a}(t)}{w_{H,a_0}(t)}\right) &= d_{H,a,a_0} - \frac{1}{\tau^H} \log\left(\frac{H_a(t)}{H_{a_0}(t)}\right) + u_{H,a,a_0} \\ \log\left(\frac{w_{L,a}(t)}{w_{L,a_0}(t)}\right) &= d_{L,a,a_0} - \frac{1}{\tau^L} \log\left(\frac{L_a(t)}{L_{a_0}(t)}\right) + u_{L,a,a_0} \end{aligned} \tag{7}$$

Where a_0 is the reference age category. Age productivities $\lambda_{H,a}$, $\lambda_{L,a}$ can then be retrieved from fixed effect d_{H,a,a_0}, d_{L,a,a_0} using normalization conditions (3).

Estimates for aggregate labor inputs $H(t)$ and $L(t)$ can be computed from estimated age productivities and elasticities of substitution. First order conditions (5) also give an expression for relative wage across education levels:

$$\log\left(\frac{w_{H,a}(t)}{w_{L,a}(t)}\right) - \log\left(\frac{\left(\lambda_{H,a}^{\tau_H-1}\frac{H(t)}{H_a(t)}\right)^{\frac{1}{\tau^H}}}{\left(\lambda_{L,a}^{\tau_L-1}\frac{L(t)}{L_a(t)}\right)^{\frac{1}{\tau^L}}}\right) = \frac{\sigma-1}{\sigma}\log\left(\frac{\theta_H(t)}{\theta_L(t)}\right) - \frac{1}{\sigma}\log\left(\frac{H(t)}{L(t)}\right) \tag{8}$$

Assume $\log\left(\frac{\theta_H(t)}{\theta_L(t)}\right)$ follows a linear time trend. Plugging in previously estimated age productivities and elasticities of substitution and adding measurement error gives us equation

$$\log\left(\frac{w_{H,a}(t)}{w_{L,a}(t)}\right) - \hat{f} = l(t) - \frac{1}{\sigma}\log\left(\frac{H(t)}{L(t)}\right) + v_{a,t} \tag{9}$$

Where $l(t)$ is a linear function of time and $\hat{f}$ is estimated from equations (7).

Weighted Least Square estimation of equations (7) and (9) are presented in table 1 and 2. The weights used are the inverse sampling variance of estimated wage gaps[2]. Labor input from any given education level and age bin is computed as the total sum of hours workers per month in a year. Average wage premia between age and within education are used as outcome variable in equation (7) and computed yearly and by education level by regressing individual wages on a dummy for age, plus fixed effects for gender, industry and region, to control for composition effects. Average wage premia between education levels and within ages are computed in the same fashion.

Table 1: Estimated age productivities and elasticities of substitution - Reduced Form

	Below High School	Above High School
τ	15.907	15.301
	(2.216)	(2.427)
λ_y	0.332	0.331
	(0)	(0.001)
λ_m	0.333	0.332
	(0)	(0)
λ_s	0.335	0.338
	(0)	(0.001)
R^2	0.994	0.972
Obs.	58	58

Estimated age elasticities of substitution τ in Portugal from 1987 to 2017 are higher than

[2]In equation (9), I weight by the inverse of the sum of the wage gaps and $\hat{f}$ inverse sampling variance

estimates found by Card and Lemieux (2001) for the US, the UK and Canada from the 1970s to the early 1990s, which are between 4 and 6. This reflects the lesser impact of movements in relative age group supply on age group wage differential in Portugal than in the US, UK and Canada. Estimated age productivities are very similar between education levels. They are also balanced between age groups, which suggests no age group is much more productive than another.

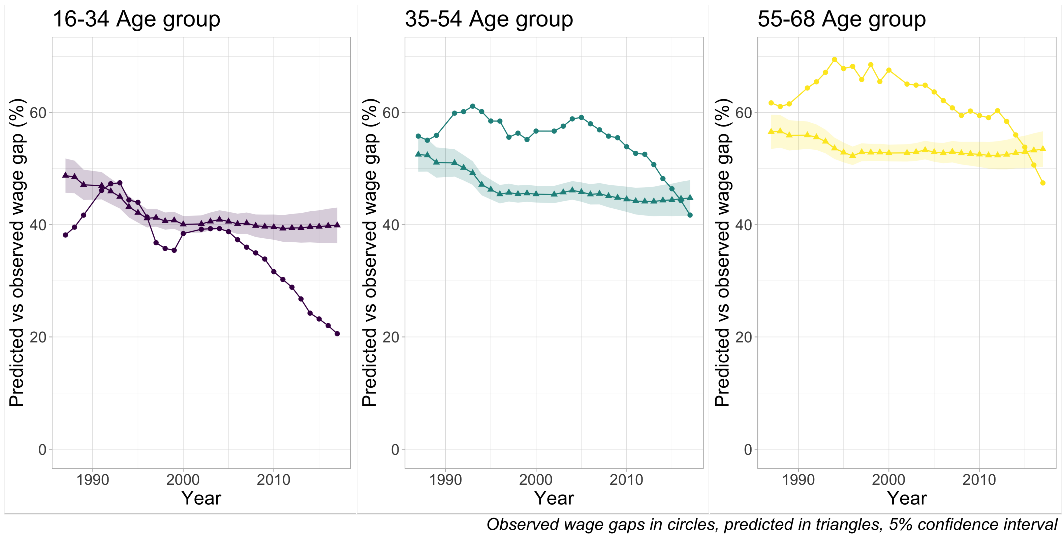
Table 2: Estimated education productivity growth and elasticity of substitution - Reduced Form

σ	4.933
	(0.151)
$\log \frac{\theta_H}{\theta_L}$	0.018
	(0.001)
R^2	0.974
Obs.	87

Elasticity of substitution between workers below and above high school is also higher in Portugal than what is found by Katz and Murphy (1992) for the US and Card and Lemieux (2001) for the UK and the US, who has estimates between 2 and 2.5. However Card and Lemieux (2001) find no significant effect of relative labor supply on relative wage between education levels in Canada, suggesting a very high substitutability of graduates and non graduates in that country. their analysis also focuses on college versus high school graduates, which is not directly comparable to my analysis on high school graduates and non graduates, who appear to be more substitutable than college graduates and non graduates. Like Katz and Murphy (1992) and Card and Lemieux (2001), I find evidence of skill-biased technological change in Portugal over the period, as relative productivity between education groups increases by 1.6% every year. This is in the range of what Card and Lemieux (2001) find for the US, UK and Canada.

This analysis informs on the large substitutability of workers between age groups and education levels, as well as the slow but significant high school biased technological change occurring in the Portuguese economy between 1987 and 2017, under simple assumptions on supply and demand. Its conclusion is that it is the increase in relative supply of high school graduates that causes the decrease in wage premium, in particular for young workers, who experience a more important rise in relative supply. Due to the high elasticities of substitution however, the effect on wage premia of a rise in relative supply is expected to be modest. Indeed, figure 3 presents the predicted wage gaps by age group. If it matches relatively well observed wage gaps for senior and middle-aged workers, it fails at reproducing the observed drop in young workers' wage premium.

Figure 3: Predicted wage gaps between high school graduates and non-graduates of same age



In order to go beyond the limitations of this simple supply and demand model, I introduce in the next section a model of matching that accounts for worker preferences, sorting, and varying worker substitutability over time.

5 Model

Recent administrative matched employer-employee datasets such as the one I use for Portugal hold much more information than workers' characteristics and wage. They also inform on firms attributes and one the joint distribution of workers and firms, that is, who is hired by who. Besides matching, we also observe the transfers mage between agents, in the form of wage. using these types of datasets enables to build a much richer framework to understand the race between education and technology, that encompasses both single firms profit maximization and workers' preferences for different types of firms. Worker productivity can vary depending on the type of firm they work for, and the extent of the impact of an increase in relative supply can vary over time. In order to take full advantage of the dataset, I build a model of one-to-many matching with transferable utility. It is akin to Azevedo and Hatfield (2018)'s work on competitive equilibrium in an economy where buyers and sellers match on a finite trades, with no limit on how many of them can enter the trade. I restrict their framework by giving more structure to the match, that several workers but only a single firm can enter. I allow firms and workers to choose from an infinite set of "trades" (in my case workforces). Existence of equilibrium rests on a large market assumption, as in Azevedo and Hatfield (2018) and Che et al. (2019). In order to bring the model to the data, I borrow to Galichon and Salanie (2020) by modelling unobserved heterogeneity in the form of random utility. This enables me to make predictions on equilibrium matching and wages from my model.

5.1 Set Up

The labor market is two-sided, with workers and firms on each side. There is a continuum of workers $i \in I$. Each worker has a type $x \in \mathcal{X}$. Types are discrete and possibly multidimensional. There is a mass n_x of workers of type x, and a finite number of types: $\#\mathcal{X} = N$. On the other side of the market, there is a large number of firms $j \in J$. Each firm has a type $y \in \mathcal{Y}$. Types are discrete and possibly multidimensional. There is a mass m_y of firms of type y, and a finite number of types: $\#\mathcal{Y} = M$.

Each firm matches with a non negative mass of worker of each type. Let k_x be the mass of type x workers a firm is matched with. Vector k is the workforce employed by the firm, where $k = (k_1, \dots, k_N) \in \mathbb{R}^N$ Each worker matches with a single firm.

Type x worker's utility for being employed at type y firm within workforce k is $u_{xy}(k)$. It is assumed to be additive in a level of amenity α that depends both on worker and firm type, as well as wage w paid by the firm to the worker:

$$u_{xy}(k) = \alpha_{xy}(k) + w_{xy}(k) \tag{10}$$

Similarly, firm profit is additive in production and total wage paid to its workforce:

$$v_y(k) = \gamma_y(k) - \sum_{x=1}^{N} k_x w_{xy}(k) \tag{11}$$

I restrict worker amenities functions $(\alpha_{xy})_{x\in\mathcal{X},y\in\mathcal{Y}}$ and firm productivities functions $(\gamma_y)_{y\in\mathcal{Y}}$ to the set of L^2 functions.

Assumption 1. *$\alpha_{xy}, \gamma_y \in L^2$, for all x, y:*

$$\int |\alpha_{xy}(k)|^2 \, dk < \infty \text{ and } \int |\gamma_y(k)|^2 \, dk < \infty$$

where notation $\int dk$ is short for $\int_0^{n_1} \dots \int_0^{n_N} dk_N \dots dk_1$

A market is characterized by exogenous distributions of worker and firm types $(n_x)_{x\in\mathcal{X}}$ and $(m_y)_{y\in\mathcal{Y}}$, as well as amenity functions $(\alpha_{xy})_{x\in\mathcal{X},y\in\mathcal{Y}}$ and productivity functions $(\gamma_y)_{y\in\mathcal{Y}}$. I describe the competitive equilibrium in the next section.

5.2 Competitive Equilibrium

Loosely speaking, competitive equilibrium on any market is found when supply from workers meets demand from firms. Supply and demand should also be feasible: workers should not ask for more firms y than m_y, nor should firms ask for more than n_x workers of type x. Formally, supply and demand are defined below.

Definition 1. *Supply*
Define supply $S = (S^x)_{x\in\mathcal{X}}$ where $S^x_y(k)$ is the mass of type x workers willing to match with type y firm and mass k workforce and $S^x(\varnothing)$ is the mass of type x workers willing to remain unmatched. S is feasible iff $S \in \mathcal{S}$ where[3]

$$\mathcal{S} = \left\{ (S^x)_x \,,\, \middle| \sum_y \int S^x_y(k)dk + S^x(\varnothing) = n_x \,\forall x \text{ and } \int \frac{S^x_y(k)}{k_x} dk \le m_y \,\forall x \in \mathcal{X},\, \forall y \right\}$$

Set $\mathcal{S}$ describes the set of feasible supplies: it imposes first that all worker mass n_x is accounted for, and second that over all workforces k, the number of type y firms that type x workers are willing to supply does not exceed m_y.

Definition 2. *Demand*
Define demand $D = (D^y)_{y\in\mathcal{Y}}$ where $D^y(k)$ is the mass of type y firms willing to match with mass k workforce and $D^y(\varnothing)$ is the mass of type y firms willing to remain unmatched. D is feasible iff $D \in \mathcal{D}$ where

$$\mathcal{D} = \left\{ (D^y)_y \,,\, \middle| \int D^y(k)dk + D^y(\varnothing) = m_y \,\forall y \text{ and } \sum_y \int k_x D^y(k)dk \le n_x \,\forall y \right\}$$

Set $\mathcal{D}$ describes the set of feasible demands: all m_y firms must be accounted for, and the sum of type x mass demanded by all firms over all workforces must not exceed the total mass of type x workers available on the market n_x.

The difference between supply and demand at the worker type, firm type, and workforce level is defined as excess demand.

[3] To avoid integrability issues in feasibility condition, consider the function defined as $\frac{S^x_y(k)}{k_x}$ if $k_x > 0$ and 0 if $k_x = 0$

Definition 3. *Excess Demand*
Given types x, y and workforce mass k, excess demand is defined as

$$Z_{xyk}(S, D) = k_x D^y(k) - S_y^x(k)$$

In the context of one-to-many matching, supply S and demand D are measured in different 'units': if a firm can match with numerous workers types, workers can only match with one firm type. This shows in the feasibility conditions for S and D: a mass $S_y^x(k)$ of type x workers supplying firms of type y within workforce k corresponds to $\frac{S_y^x(k)}{k_x}$ firms supplied. However a mass $D^y(k)$ demanding k_x workers corresponds to $k_x D^y(k)$ workers demanded.

A competitive equilibrium is reachedon the market when supply and demand are feasible, excess demand is zero, and matching is incentive compatible, in the following sense:

Definition 4. *Competitive Equilibrium*
An arrangement $[S, D, w]$ is a competitive equilibrium if it satisfies:

(i) S and D are feasible: $S \in \mathcal{S}$ and $D \in \mathcal{D}$

(ii) S and D clear the market: $\forall x \in X$, $y \in Y$, $k \in \mathcal{K}$: $Z_{xyk}(S, D) = 0$

(iii) Given w, each agent obtains their optimal utility/production:

$$\forall i \in I, S_{y^*}^{x_i}(k^*) > 0 \Rightarrow (y^*, k^*) \in \arg\max_{y,k} u_{x_i y}(k)$$
$$\forall j \in J, D^{y_j}(k^*) > 0 \Rightarrow k^* \in \arg\max_{k} v_{y_j}(k)$$

Wages w are transferable utility that clear the market and ensure incentive compatibility, in the sense that all agents obtain de best possible utility or profit available on the market, given wages.

The existence a competitive equilibrium on any market is not trivial. To show there always exists one in my framework, I follow a proof technique introduced in the continuum assignment problem by Gretsky et al. (1992) and Gretsky et al. (1999). The reasoning is also very close to Azevedo and Hatfield (2018)'s proof for competitive equilibrium existence in a large economy on a market of buyers and sellers with a finite set of possible trades. I expand on their proof by considering a continuous set of workforces. Define the social planner problem as the supremum of total surplus under the constraint that excess demand is zero

$$\sup_{S\in\mathcal{S},D\in\mathcal{D}} \sum_{x,y}\int S_y^x(k)\alpha_{xy}(k)\,dk + \sum_y \int D^y(k)\gamma_y(k)\,dk$$
$$\text{s.t } Z_{xyk}(S,D)=0 \quad \forall x,y,k \tag{SP}$$

Then the following holds

Theorem 1. *The maximum is attained in* (SP)*, and the optimal* (S^*, D^*) *is part of a competitive equilibrium*

Proof. The proof is in two steps: to show that the maximum is attained in (SP) show that the objective function is continuous on a compact set. Then to see that the optimal (S^*, D^*) is part of a competitive equilibrium, use complementary slackness conditions to show that the dual variables to (SP) at optimum are incentive compatible. The full proof can be found in appendix B □

Theorem 1 provides a proof of existence for the competitive equilibrium, but also a practical way of computing it, through solving problem (SP). In order to bring the model to the data, I add one last dimension to the model that accounts for unobserved variables in match formation.

5.3 Matching heterogeneity

Some firm and workers characteristics that play a role in match formation are unobserved, and therefore are not accounted for in x or y. There exists a large literature that deals with unobserved heterogeneity, and I build on a large subset (**?**, **?**) that uses additive random shocks to model it. I further assume a logit framework for the model by restraining the distribution of shocks to belong the extreme value class, although as shown in **?** in the one-to-one case identification is possible with a general class of distributions.

Workers i experiences stochastic shock $(\epsilon_{iy}(k))_{y,k}$ in addition to her systematic utility:

$$u_{x_iy}(k) + \epsilon_{iy}(k) \tag{12}$$

Similarly firm j experiences stochastic shock $\eta_j(k)$ in addition to its systematic production:

$$v_{y_j}(k) + \eta_j(k) \tag{13}$$

I impose the following independence conditions on stochastic shocks:

Assumption 2. *Stochastic shocks satisfy the following:*

(i) For each pair of two workers i and i', $\epsilon_{iy}(k)$ and $\epsilon_{i'y}(k)$ are mutually independent and identically distributed

(ii) For each pair of two firms j and j', $\eta_j(k)$ and $\eta_{j'}(k)$are mutually independent and identically distributed

(iii) For a worker i and a firm j $\epsilon_{iy}(k)$ and $\eta_j(k)$ are mutually independent

(iV) $\epsilon_{iy}(k)$ is independent of α_{x_iy}, $\eta_j(k)$ is independent of γ_y

I further restrict stochastic shocks to the analog of logit shocks in a continuous framework:

Assumption 3.

$$\epsilon_{iy}(k) = \max_n \{\epsilon_i^n : k_i^n = k, y_i^n = y\} \text{ and } \eta_j(k) = \max_n \left\{\eta_j^n : k_j^n = k\right\}$$

Where $(y^n, k^n, \epsilon_i^n)_n$, $(k^n, \eta_j^n)_n$ are enumerations of two inhomogeneous Poisson point processes on of intensity $e^{-\epsilon}$ and $e^{-\eta}$ respectively.

Definition 5. *Expected Utilities*
Define expected utilities on both sides of the market by

$$G_x(u) = \mathbb{E}\left[\max_{y,k}\{u_{xy}(k) + \epsilon_y(k)\}\right] \text{ and } H_y(v) = \mathbb{E}\left[\max_k\{v_y(k) + \eta(k)\}\right]$$

Under assumption 3, expected utilities rewrite in closed form.

Proposition 1. *If ϵ and η are part of Poisson point processes, then*

$$G_x(u_x) = \log \sum_y \int \exp(u_{xy}(k))\, dk \text{ and } H_y(v_y) = \log \int \exp(v_y(k))\, dk$$

Proof. In appendix B $\square$

Under assumptions 2, 3 and following Dupuy and Galichon (2014) and Galichon and Salanie (2020), I can rewrite the social planner problem with heterogeneity as the original problem, with a penalty (also called the entropy) that accounts for heterogeneity. Predictions for optimal matching and wages can also be obtained.

Theorem 2. *The social planner problem with heterogeneity*

$$\max_D \sum_y \int \Phi_y(k) D^y(k) dk + \mathcal{E}(D, n, m) \tag{SP$_{ent}$}$$

Where $\mathcal{E}(D,n,m)$ is the entropy and is equal to

$$\mathcal{E}(D,n,m) = -\sum_x n_x \sum_y \int \frac{D^y(k)k_x}{n_x} \log \frac{D^y(k)k_x}{n_x}\, dk - \sum_y m_y \int \frac{D^y(k)}{m_y} \log \frac{D^y(k)}{m_y}\, dk$$

Optimal D solves

$$\log D^y(k) = \frac{\Phi_y(k) - \sum_x k_x U_x - V_y + \sum_x k_x \log \frac{n_x}{k_x} + \log m_y}{1 + \sum_x k_x} \tag{14}$$

Market clearing wages solve

$$w_{xy}(k) = \frac{\gamma_y(k) - \alpha_{xy}(k) - V_y + U_x + \log m_y - \log \frac{n_x}{k_x}}{1 + \sum_x k_x} \tag{15}$$

Where U_x, V_y solve

$$\begin{cases} \sum_{y,k} k_x \exp\left(\frac{\Phi_{yk} - \sum_x k_x U_x - V_y + \sum_x k_x \log\left(\frac{k_x}{n_x}\right) + \log m_y}{1+\sum_x k_x}\right) = n_x \\ \sum_k \exp\left(\frac{\Phi_{yk} - \sum_x k_x U_x - V_y + \sum_x k_x \log\left(\frac{k_x}{n_x}\right) + \log m_y}{1+\sum_x k_x}\right) = m_y \end{cases} \tag{16}$$

Proof. In appendix B □

5.4 Sensitivity analysis

Theorem 2 makes predictions on both wage and matching, and allows to predict how they vary when a given parameter changes. In particular, equation (15) shows there is a direct and an indirect effect of any parameter change on wage levels: the direct effect shows directly through the parameter, while the indirect effect comes at play through the parameter's influence on expected utilities U and V. For instance, if the mass of type x workers n_x is augmented by some small amount, the effect on wage of type x workers in any given firm type y and workforce k is

$$\frac{\partial w_{xy}(k)}{\partial n_x} = \frac{-\frac{\partial V_y}{\partial n_x} + \frac{\partial U_x}{\partial n_x} - \frac{1}{n_x}}{1 + \sum_x k_x} \tag{17}$$

Where $-\frac{1}{n_x}$ is the direct effect and - $\frac{\partial V_y}{\partial n_x} + \frac{\partial U_x}{\partial n_x}$ the indirect effect. Because $n_x > 0$, the direct effect on wage is always negative. Intuitively, the indirect effect could go both ways: because U_x is the Lagrange multiplier on the marginal constraint for type x workers, and increasing n_x relaxes this constraint, $\frac{\partial U_x}{\partial n_x}$ is below 0. However the effect of n_x on V_y is ambiguous.

Other worker types x' wages are also affected by a change in n_x:

$$\frac{\partial w_{x'y}(k)}{\partial n_x} = \frac{-\frac{\partial V_y}{\partial n_x} + \frac{\partial U_{x'}}{\partial n_x}}{1 + \sum_x k_x} \tag{18}$$

Only an indirect effect comes into play. The effect of n_x on $U_{x'}$ is also ambiguous.

$\frac{\partial U_x}{\partial n_x}$, $\frac{\partial U_{x'}}{\partial n_x}$ and $\frac{\partial V_y}{\partial n_x}$ can be computed numerically by linearization of equations (16).

6 Identification and Estimation

The model's predictions on matching (14) and wage (15) allow to separately identify amenity and productivity functions $(\alpha_{xy})_{xy}$ and $(\gamma_y)_y$. This would not be true if we observed only matching, as α and γ appear together in the matching prediction, and only total surplus can be identified from this equation. If only wages were observed, the same problem arises and only the difference between firm production and worker amenities is identified. One must assume that amenities are zero in order to identify production, as in the reduced forms in section 4.

I am aiming to parametrically estimate α and γ. The functional forms are the same as the ones laid out in section 3: there are $N = 6$ worker types that are the combination of two education levels, high school graduates and non graduates, and three age groups, young (below 35), middle-aged (between 35 and 54), and senior (above 55). Let $e(x)$, $a(x)$ be type x's education level and age group. Firm workforce $k(t)$ at time t is composed of employed masses of each worker types:

$$k(t) = (k_{H,y}(t), k_{H,m}(t), k_{H,s}(t), k_{L,y}(t), k_{L,m}(t), k_{L,s}(t))$$

Employed mass of worker k_x is defined as total number of hours worked monthly by workers of type x hired by the firm, divided by 174, the monthly hours equivalent of a 40 hours week. Hence each k_x counts the full-time equivalent of the number of type x workers employed by the firm. This measure is not necessarily an integer, as part-time workers would count as fractions of the full-time equivalent. I allow all amenity and production parameters to vary over time. and Type y firm production $\gamma_y^t(k)$ at time t is

$$\gamma_y^t(k) = \left[(\theta_H^y(t) H(t))^{\frac{\sigma(t)-1}{\sigma(t)}} + (\theta_L^y(t) L(t))^{\frac{\sigma(t)-1}{\sigma(t)}} \right]^{\frac{\sigma(t)}{\sigma(t)-1}}$$

Where aggregates $H(t)$ and $L(t)$ are:

$$\begin{aligned} H(t) &= \left[\sum_{a \in \{y,m,s\}} \lambda_{a,H}(t) k_{H,a}(t)^{\frac{\tau^H(t)-1}{\tau^H(t)}} \right]^{\frac{\tau^H(t)}{\tau^H(t)-1}} \\ L(t) &= \left[\sum_{a \in \{y,m,s\}} \lambda_{a,L}(t) k_{L,a}(t)^{\frac{\tau^L(t)-1}{\tau^L(t)}} \right]^{\frac{\tau^L(t)}{\tau^L(t)-1}} \end{aligned} \tag{19}$$

I assume worker amenities are constant in k:

$$\alpha^t_{xy}(k) = \beta^y_x(t) \tag{20}$$

β^y_x reflects type x worker preferences for type y firms over other firm types. In particular I assume workers are indifferent to workforce size.

Given these functional forms, I am looking to estimate in every period t parameters $\left(\lambda^y_{a,H}(t)\right)_a$, $\left(\lambda^y_{a,L}(t)\right)_a$, $(\theta^y_H(t))_y$, $(\theta^y_L(t))_y$, $(\beta^y_x(t))_{x,y}$, $\tau_H(t)$, $\tau_L(t)$ and σ. The first step estimates amenities parameters and the second production parameters.

Step 1: I estimate amenity parameters $(\beta^y_x(t))_{x,y}$ through equation (15). Assume $\tilde{w}_{ij\kappa}$, worker i's wage for working at firm j within workforce κ is observed in the data with measurement error, i.e. for any individual worker i of type x employed firm j of type y within workforce κ of size k:

$$\tilde{w}_{ij\kappa} = w_{xy}(k) + \epsilon_{ij\kappa} \text{ where } \epsilon \sim \mathcal{N}(0, s^2) \text{ iid} \tag{21}$$

Given the previous assumption, average wage paid to type x workers in firm j that employs workforce k is normally distributed with mean 0 and variance $k_x s^2$.

Let x^* be any reference worker type. Then differentiating equation (15) for a given worker type x, workforce k and industry y obtains:

$$w_{x^*y}(k) - w_{xy}(k) = \frac{1}{1 + \sum_x k_x} \left(U_{x^*} - U_x + \beta^y_x - \beta^y_{x^*} + \log\left(\frac{n_x k_{x^*}}{n_{x^*} k_x} \right) \right) \tag{22}$$

Where production $\gamma_y(k)$ has been differentiated out.

In order to use equation (23) to estimate β^y_x, I choose reference type x^* to be non high school graduates between 16 and 34. and I make the following assumption:

Assumption 4. *Reference type x^* if indifferent between firm types, and has null amenity parameters:* $\beta_{x^*}^{y} = 0$

Under the previous assumption, equation (23) at firm j's level becomes

$$\bar{w}_{x^*j\kappa} - \bar{w}_{xj\kappa} - \log\left(\frac{n_x k(\kappa)_{x^*}}{n_{x^*} k(\kappa)_x}\right) = \frac{1}{1 + \sum_x k(\kappa)_x}\left(U_{x^*} - U_x + \beta_x^{y_j}\right) + \epsilon_{j\kappa}$$
$$\text{where } \epsilon_{j\kappa} \sim \mathcal{N}(0, (k(\kappa)_{x^*} - k(\kappa)_x)s^2) \text{ iid} \tag{23}$$

Where $\bar{w}_{x^*j\kappa}$,$\bar{w}_{xj\kappa}$ are average type x^* and x wages paid by firm j, and $k(\kappa)_x$ is the mass of type x workers employed in workforce κ. Hence amenity parameters can be estimated with Generalized Least Squares through

$$y_{j\kappa x} = a_x K(\kappa) + b_{xy} K(\kappa) + \epsilon_{j\kappa} \tag{24}$$

Where $y_{j\kappa x} = \bar{w}_{x^*j\kappa} - \bar{w}_{xj\kappa} - \log\left(\frac{n_x k(\kappa)_{x^*}}{n_{x^*} k(\kappa)_x}\right)$ and $K(\kappa) = \frac{1}{1+\sum_x k(\kappa)_x}$. b_{xy} is identified because of variation in $K(\kappa)$. There is however some colinearity in the equation, due to the presence of a_x, which leads to the following additional assumption:

Assumption 5. $\beta_x^{y^*} = 0$ *for some firm type* y^*

I choose y^* to be the Agriculture, Mining, Energy and Construction industry.

Step 2: In a second step, I estimate remaining parameters $(\lambda_{a,H}(t))_a$, $(\lambda_{a,L}(t))_a$, σ, τ_h, τ_l through maximum likelihood on matching. Let $\tilde{\mu}_{yk}$ be the observed mass of type y firms matched with type k workforce. Let μ_{yk} be the prediction (14) made by the model. In particular μ_{yk} satisfies margin equations (16). The log likelihood is:

$$\sum_{y,k} \tilde{\mu}_{yk} \log \mu_{yk}$$

Where factor $(1 + \sum_x k_x)$ weighs the likelihood by the number of individual observed. Maximizing the log likelihood requires to solve

$$\max_{\sigma,\tau_l,\tau_h} \sum_{y,k} \tilde{\mu}_{yk}\left(\Phi_{yk} - \sum_x k_x U_x - V_y + \sum_x k_x \log\left(\frac{k_x}{n_x}\right) + \log m_y\right)$$
$$\text{s.t } U, V \text{ statisfy (16)} \tag{LL}$$

Given values for $(\beta_x^y)_{x,y}$, productivity parameters $(\theta_h^y)_y$, $(\theta_l^y)_y$, $(\lambda_{a,H}(t))_a$, $(\lambda_{a,L}(t))_a$, and elasticities of substitution σ, τ_h, τ_l are identified through first order conditions to problem (LL).

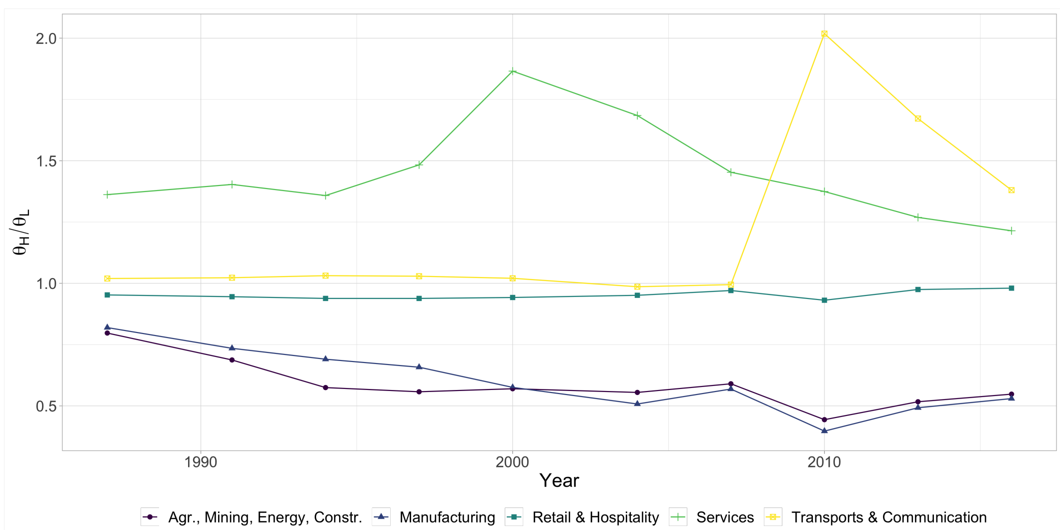
I run both estimation steps on ten separate three year periods between 1987 and 2017[4].

7 Results

7.1 Parameters estimates

Estimates for $\frac{\theta_H^y}{\theta_L^y}$ by industry y are presented in figure 4. They show evidence of differing relative demand by industry: industries for which $\frac{\theta_H^y}{\theta_L^y}$ is below 1, such as manufacturing and agriculture, mining, energy, construction, have a higher demand for non high school graduates than graduates. Other industries such as services and transport and communications have a higher demand in favor of graduates. Relative demand does not evolve in the same way depending on industry either: it is increasingly biased towards non graduates in manufacturing and agriculture, mining, energy, construction, and constant in retail and hospitality and transport and communications until 2007. Services displays a particular evolution for demand: it starts increasing in favor of high school graduates until 2000-2003, but then changes direction in favor of non graduates over the rest of the period.

Figure 4: Estimated education productivities

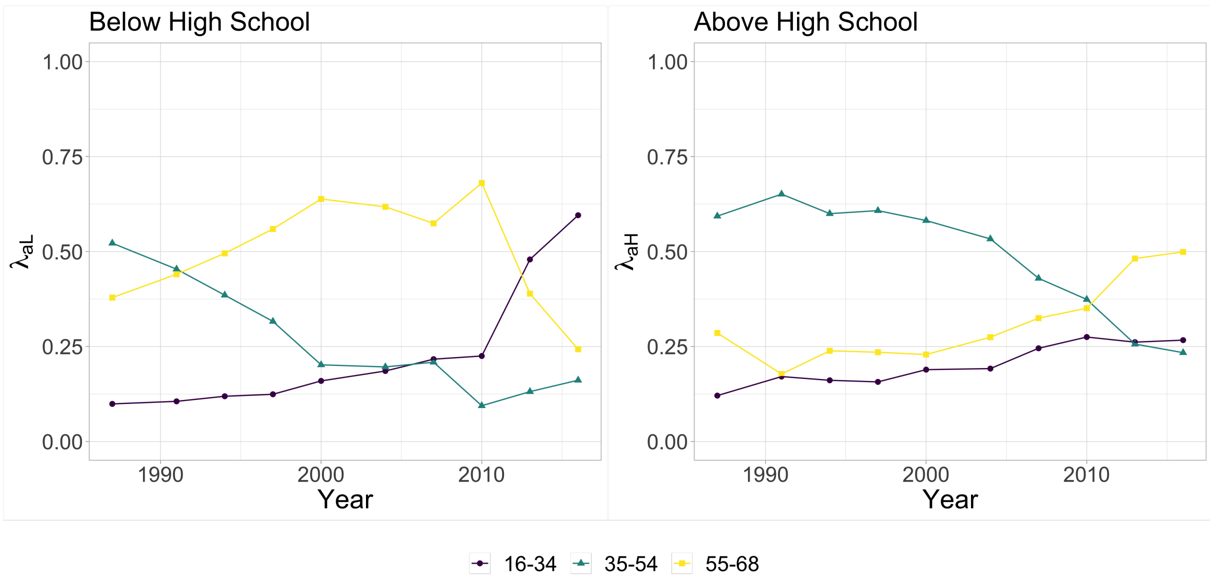


Graph 5 shows the evolution of age productivities by education level $\lambda_{H,a}$ and $\lambda_{L,a}$ by age

[4]Since data for years 1990 and 2001 are missing, the last time period spans only two years.

group and over time. young workers' productivity increases and middle-aged workers' decreases for both education levels However senior workers' productivity decreases for non high school graduates and increases for high school graduates.

Figure 5: Estimated age productivities



Graph 6 presents the change in worker preferences for firms β_x^yn in euros per hours worked. It is computed with respect to reference industry agriculture, mining, energy, construction, and worker type Below high school, between 16 and 34. It shows most worker types experience preferences that are close to the reference worker type, except for senior high school graduates, whose amenities for working are decreasing in all industries over the period, especially in manufacturing.

Figure 6: Estimated worker preferences

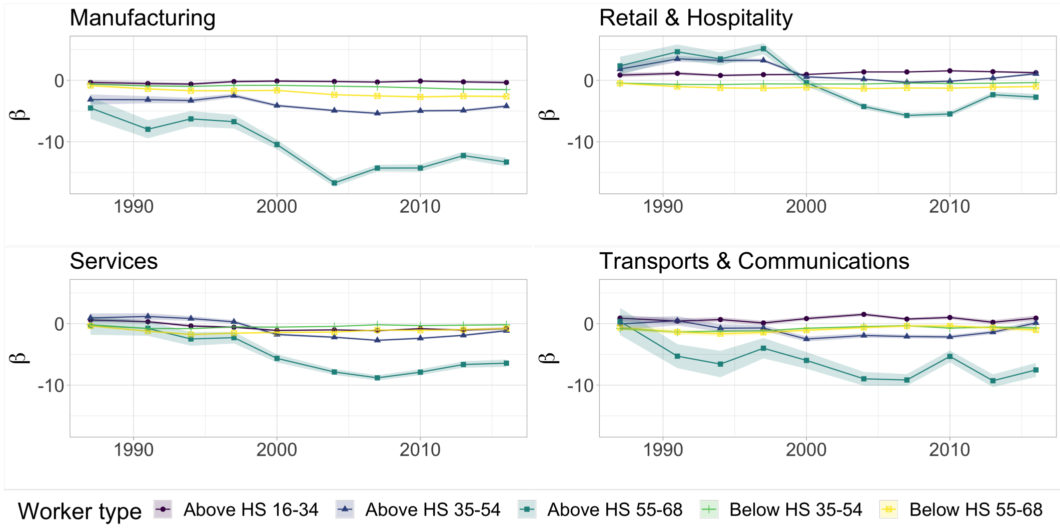
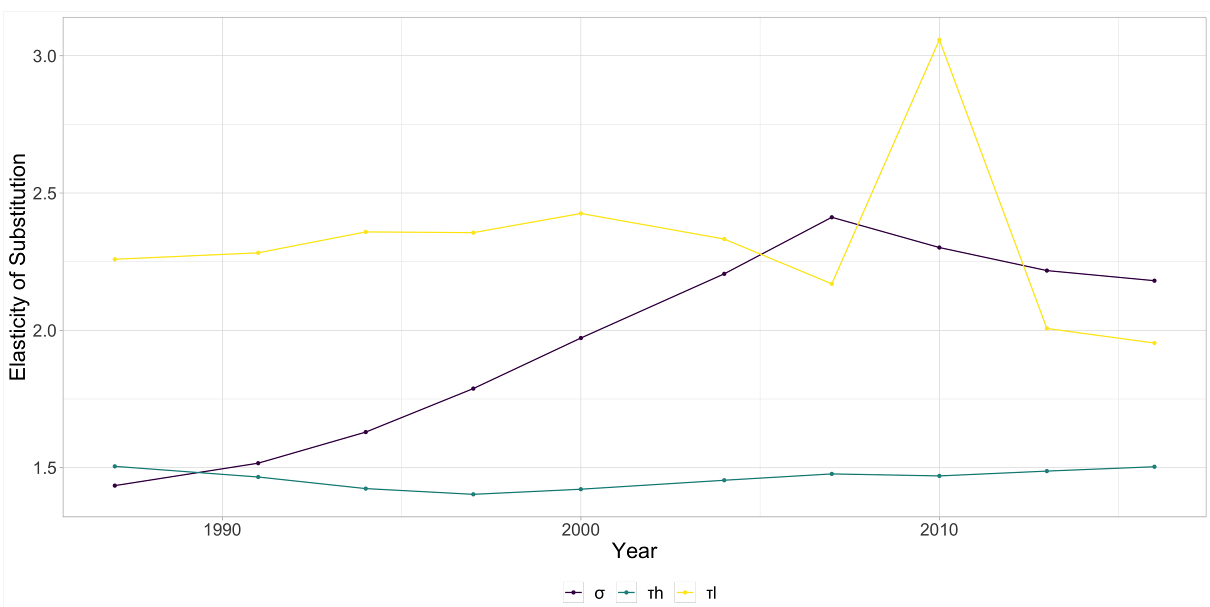


Figure 7 presents estimated elasticities of substitution[5]. It shows substitution between age groups is higher among non high school graduates than graduates. Substitution between education levels has increased steadily until 2007, and stagnated since.

Figure 7: Estimated elasticities of substitution



[5]This a temporary estimates and are subject to change, as I explore the log likelihood maximization space more thoroughly

References

Autor, D., C. Goldin, and L. F. Katz
2020. Extending the Race between Education and Technology. *AEA Papers and Proceedings*, 110:347–351.

Autor, D. H., F. Levy, and R. J. Murnane
2003. The Skill Content of Recent Technological Change: An Empirical Exploration*. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Azevedo, E. M. and J. W. Hatfield
2018. Existence of Equilibrium in Large Matching Markets With Complementarities. SSRN Scholarly Paper ID 3268884, Social Science Research Network, Rochester, NY.

Beaudry, P. and D. Green
2005. Changes in U.S. Wages, 1976–2000: Ongoing Skill Bias or Major Technological Change? *Journal of Labor Economics*, 23(3):609–648. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Beaudry, P., D. A. Green, and B. M. Sand
2015. The Great Reversal in the Demand for Skill and Cognitive Tasks. *Journal of Labor Economics*, 34(S1):S199–S247. Publisher: The University of Chicago Press.

Blair, P. Q. and D. J. Deming
2020. Structural Increases in Demand for Skill after the Great Recession. *AEA Papers and Proceedings*, 110:362–365.

Card, D. and J. E. DiNardo
2002. Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. Working Paper 8769, National Bureau of Economic Research. Series: Working Paper Series.

Card, D. and T. Lemieux
2001. Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics*, 116(2):705–746.

Che, Y.-K., J. Kim, and F. Kojima
2019. Stable Matching in Large Economies. *Econometrica*, 87(1):65–110. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA13547.

Dupuy, A. and A. Galichon
2014. Personality Traits and the Marriage Market. *Journal of Political Economy*, 122(6):1271–1319.

Ekeland, I., J. Heckman, and L. Nesheim
2004. Identification and Estimation of Hedonic Models. *Journal of Political Economy*, 112(S1):S60–S109. Publisher: The University of Chicago Press.

Galichon, A.
2016. *Optimal Transport Methods in Economics*. Princeton University Press.

Galichon, A. and B. Salanie
2020. Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *SSRN Electronic Journal*.

Goldin, C. and L. Katz
2008. *The Race Between Education and Technology*. Belknap Press for Harvard University Press.

Gretsky, N., J. Ostroy, and W. Zame
1992. The Nonatomic Assignment Model. *Economic Theory*, 2:103–27.

Gretsky, N. E., J. M. Ostroy, and W. R. Zame
1999. Perfect Competition in the Continuous Assignment Model. *Journal of Economic Theory*, 88(1):60–118.

Hatfield, J. W. and P. R. Milgrom
2005. Matching with Contracts. *American Economic Review*, 95(4):913–935.

Hershbein, B. and L. B. Kahn
2018. Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings. *American Economic Review*, 108(7):1737–1772.

Katz, L. F. and K. M. Murphy
1992. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Kelso, A. S. and V. P. Crawford
1982. Job Matching, Coalition Formation, and Gross Substitutes. *Econometrica*, 50(6):1483–1504.

Tinbergen, J.
1974. Substitution of Graduate bu Other Labour.

Valletta, R. G.
2016. Recent Flattening in the Higher Education Wage Premium: Polarization, Skill Downgrading, or Both? Working Paper 22935, National Bureau of Economic Research. Series: Working Paper Series.

A Data

I use *Quadros de Pessoal*, a matched employer-employee dataset provided by the Portuguese National Institute (Instituto Nacional de Estatística, INE). *Quadros de Pessoal* is issued yearly from 1987 to 2017, based on firms declarations on their characteristics and their employees'. Both workers and firms are identified across time by a unique identifier.

I use information on firm industry, worker's age and education level Industries are provided as "economic activity", up to 3 digit level. Because of classification changes at the 2 and 3 digits level over time, I use the one digit level classification, to keep consistency over the years. I exclude firms whose economic activity at the 1 digit level are unknown. Worker education is provided as a 3 digits classification, out of which I aggregate 9 levels: no schooling, primary schooling 1 (up to 10 years old), primary schooling 2 (up to 13 years old), primary schooling 3 (up to 15 years old), completed high school, some higher education, bachelor, masters and PhD. Worker age is used directly without further cleaning. I exclude from the sample any worker whose education level of age is unknown (3.9% of observations per year on average)

I also use information on wages and number of hours worked. Wage is provided as a average monthly earnings, that accounts for bonuses and extra hours earnings. Number of hours is provided as the baseline number of hours in the contract, plus any extra hours worked (averaged overt he year). I consider the sum of base and extra hours as my measure for number of hours worked per month. I divide monthly wage by monthly hours to obtain a measure of hourly wage, and deflate it. Real hourly wage is my final measure of wage. I exclude from the sample any worker who has worked zero hours or earned zero wage over the year (11.5% of observations per year on average). These are mainly, in my understanding, workers on sick leave, maternity leave, or sabbatical that do not contribute to firm production in that year. I also exclude from the sample any workers who are strictly under 16 or above 68 (the retirement age in Portugal)

Additionally, I exclude any observation with a missing or 0 worker ID (3.5% of observations per year on average). I am also faced with an issue of duplicate worker IDs which, even though it is minor in the sample later years (about 4.8% of observations per year on average from 2007 to 2017, including 0 IDs), it is much more serious in the earlier years (about 19% of the sample in 1987, including 0 IDs). I suspect these to be encoding mistakes that relate to actual different workers. Some can also be workers who hold two different jobs (for instance an employee somewhere who also have a self-employed activity). Because I do not use the panel aspect of the data, and therefore encoding mistakes in workers ID are not a problem in my analysis, I keep most duplicates, only removing observations who appear more than 5 times in any given year (an average 6.1% of observations per year, less than 1% of the dataset starting in 2007). I also exclude from the sample any worker who is self-unemployed, in unpaid family care, or labelled under "other" employment contract (7.1% of observations per year on average). The rationale behind not considering self-employed is that many of self-employed workers actually work as consultants for a firm, with no way to link them. Self-employed workers on their own represent about 1% of the dataset.

B Proofs

Theorem 1

Proof. **Step 1**: Show that the maximum is attained in (SP):

(i) Function $W(S,D) = \sum_{x,y} \int S^x_y(k)\alpha_{xy}(k)\,dk + \sum_y \int D_y(k)\gamma_y(k)\,dk$ is continuous

(ii) $\mathcal{L} = \{(S,D)\,|S^x \in \mathcal{S}, D^y \in \mathcal{D},\, Z_{xyk}(S,D) = 0\,\forall x,y,k\}$ is compact

To show (i): let sequence $(S_n, D_n)_n$ such that $(S_n, D_n) \to (S^*, D^*)$. Also let $f_n(k) = \sum_{x,y} S^x_n(y,k)\alpha_{xy}$ and $g_n(k) = \sum_y D^y_n(k)\gamma_y(k)$.

Then $f_n \to f^*$ and $g_n \to g^*$ pointwise where $f^*(k) = \sum_{x,y} S^{x,*}(y,k)\alpha_{xy}(k)$ and $g^*(k) = \sum_y D^{y,*}(k)\gamma_y(k)$.

Besides $|f_n(k)| \leq |\sum_{x,y} n_x\alpha_{xy}|$ and $|g_n(k)| \leq |\sum_y m_y\gamma_y(k)|$. Both upper bounds are integrable by assumption 1. Hence by the dominated convergence theorem $\int f_n(k)dk \to \int f^*(k)\,dk$ and $\int g_n(k)\,dk \to \int g^*(k)\,dk$

To show (ii), first note that $\mathcal{S}$ and $\mathcal{D}$ are compact for the weak topology: consider a sequence $(S_n)_n \in \mathcal{S}$ such that $S_n \rightharpoonup S$. Fatou's lemma ensures that $\int \frac{S^x(y,k)}{k_x}\, dk \leq m_y$ so that $S \in \mathcal{S}$. Therefore any sequence in $\mathcal{S}$ has a closure point in $\mathcal{S}$ and $\mathcal{S}$ is compact. So is $\mathcal{D}$ by the same reasoning.

$\mathcal{S} \times \mathcal{D}$ is compact by Tychonoff's theorem. Hence $\mathcal{L}$ is compact, because it is a closed subet of $\mathcal{S} \times \mathcal{D}$, by continuity of Z_{xyk}

Step 2: Show that $(S^*, D^*) \in \arg\max_{(S,D)\in\mathcal{L}} W(S,D)$ is part of an equilibrium.

Define total match surplus Φ:

$$\Phi_y(k) = \gamma_y(k) + \sum_x k_x \alpha_{xy}(k)$$

Then problem (SP)'s objective is

$$\begin{aligned}
&\sum_{x,y} \int S^x_y(k)\alpha_{xy}(k)\, dk + \sum_y \int D^y(k)\gamma_y(k)\, dk \\
=&\sum_{x,y} \int k_x \int D^y(k)\alpha_{xy}(k)\, d_{k-x} dk_x + \sum_y \int D^y(k) \left(\Phi_y(k) - \sum_x k_x \alpha_{xy}(k) \right) dk \\
=&\sum_y \int D^y(k)\Phi_y(k)\, dk + \underbrace{\sum_{x,y} \int k_x \int \left(D^y(k) - D^y(k)\right) \alpha_{xy}(k)\, dk_{-x} dk_x}_{=0}
\end{aligned}$$

Where we used that $Z_{xyk}(S,D) = 0$ to go from the first to the second line. Hence problem (SP) is equivalent to dual:

$$\begin{aligned}
&\max_D \sum_y \int D^y(k)\Phi_y(k)\, dk \\
&\quad \text{s.t } \sum_y \int k_x D^y(k) dk = n_x \text{ and } \int D^y(k) dk = m_y \\
\Leftrightarrow &\max_D \min_{U,V} \sum_x n_x U_x + \sum_y m_y V_y + \sum_y \int D^y(k) \left(\Phi_y(k) - \sum_x k_x U_x - V_y \right) \\
\Leftrightarrow &\min_{U,V} \sum_x n_x U_x + \sum_y m_y V_y \\
&\quad \text{s.t } \Phi_y(k) \leq \sum_x k_x U_x + V_y \, \forall y, k
\end{aligned}$$

Complementary slackness is:

$$\begin{cases} D^y(k) > 0 \Rightarrow \Phi_y(k) = \sum_x k_x U_x + V_y \\ D^y(k) = 0 \Rightarrow \Phi_y(k) < \sum_x k_x U_x + V_y \end{cases} \tag{CS}$$

Since $\max_{y,k} \alpha(y) - w_{wy}(k)$ and $\max_k \gamma_y(k) - \sum_x k_x w_{wy}(k)$ satisfy (CS), $V_y = \max_k \gamma_y(k) - \sum_x k_x w_{wy}(k)$ and $U_x = \max_{y,k} \alpha(y) - w_{wy}(k)$ are solutions to the dual of (SP) and the social planner's choice indeed yields a competitive equilibrium. □

Proposition 1

Proof. Let $Z_1 = \max_{y,k} \{u_{xy}(k)) + \epsilon_{iy}(k)\}$ and $Z_2 = \max_k \{v_y(k) + \eta_j(k)\}$. Then Z_1 follows a Gumbel distribution with expectation $\log \sum_y \int \exp(u_{xy}(k))dk$ and Z_2 follows a Gumbel distribution with expectation $\log \int \exp(v_y(k))dk$

$\mathbb{P}[Z_1 \leq c] = \prod_n \mathbb{P}[u_{xy^n}(k^n) + \epsilon_i^n \leq c]$ is the probability that process $(y^n, k^n, \epsilon_i^n)_n$ has no point in $\{(y, k, \epsilon) : u_{xy}(k) + \epsilon > c\}$. This is:

$$\begin{aligned} \mathbb{P}[Z_1 \leq c] &= \exp\left(-\sum_y \int_{c-u_{xy}(k)} e^{-\epsilon} dk d\epsilon\right) \\ \Rightarrow \log \mathbb{P}[Z_1 \leq c] &= -\sum_y \int e^{u_{xy}(k)-c} dk = -\exp\left(-c + \log \sum_y \int \exp(u_{xy}(k))dk\right) \end{aligned}$$

$\mathbb{P}[Z_2 \leq c] = \prod_n \mathbb{P}\left[v_y(k^n) + \eta_j^n \leq c\right]$ is the probability that process $(k^n, \eta_j^n)_n$ has no point in $\{(k, \eta) : v_y(k) + \eta > c\}$. This is:

$$\begin{aligned} \mathbb{P}[Z_2 \leq c] &= \exp\left(-\int_{c-v_y(k)} e^{-\eta} dk d\eta\right) \\ \Rightarrow \log \mathbb{P}[Z_2 \leq c] &= -\int e^{v_y(k)-c} dk = -\exp\left(-c + \log \int \exp(v_y(k))dk\right) \end{aligned}$$

□

Theorem 2 Based on Gretsky et al. (1992) and Galichon and Salanie (2020).

Proof. The dual problem at individual level writes

$$\begin{aligned} &\min_{U,V} \sum_i^I U_i + \sum_j^J V_j \\ &\text{s.t} \sum_{i \in \kappa} U_i + V_j \geq \Phi_{\kappa j} \quad \forall \kappa, j \end{aligned} \tag{D}$$

Where j is an individual firm and κ is a set of workers.
Let

$$\begin{cases} U_{xy}(k) = \min_{x_i=x}\{U_i - \epsilon_{iy}(k)\} \\ V_y(k) = \min_{y_j=y}\{V_j - \eta_j(k)\} \end{cases} \iff \begin{cases} U_i = \max_{y,k}\{U_{x_iy}(k) + \epsilon_{iy}(k)\} \\ V_j = \max_k\{V_{y_j}(k) + \eta_j(k)\} \end{cases}$$

Conditions in (D) become:

$$\begin{aligned} &\sum_{i\in\kappa} U_i + V_j \geq \Phi_{ky_j} + \sum_{i\in\kappa} \epsilon_{iy_j}(k) + \eta_j(k) \\ \Rightarrow &\sum_{i\in\kappa} \min_{x_i=x}\{U_i - \epsilon_{iy}(k)\} + \min_{y_j=y}\{V_j - \eta_j(k)\} = \Phi_y(k) \\ \Rightarrow &\sum_x k_x U_{xy}(k) + V_y(k) = \Phi_{ky} \end{aligned}$$

So that problem (D) is equivalent to:

$$\begin{aligned} &\min_{U,V} \sum_x n_x G_x(U) + \sum_y m_y H_y(V) \\ &\text{s.t } \sum_x k_x U_{xy}(k) + V_y(k) = \Phi_y(k) \quad \forall k, y \end{aligned}$$

Rewrite:

$$\begin{aligned} &\min_{U,V} \max_{\mu} \sum_x n_x G_x(U) + \sum_y m_y H_y(V) \\ &\quad + \sum_y \int D^y(k) \left(\Phi_y(k) - \sum_x k_x U_{xy}(k) - V_y(k) \right) dk \\ &= \max_D \sum_y \int D^y(k) \Phi_{ky} \, dk \\ &\quad - \sum_x n_x \max_U \left\{ \sum_y \int \frac{k_x D^y(k)}{n_x} U_{xy}(k) - \sum_x G_x(U) \, dk \right\} \\ &\quad - m_y \max_V \left\{ \int \sum_y \frac{D^y(k)}{m_y} V_y(k) - \sum_y H_y(V) \, dk \right\} \\ &= \max_D \sum_y \int D^y(k) \Phi_{ky} \, dk - \sum_{x,y} \int k_x D^y(k) \log \frac{k_x D^y(k)}{n_x} \, dk - \sum_y \int D^y(k) \log \frac{D^y(k)}{m_y} \, dk \end{aligned}$$

Where the last line is obtained through solving for Fenchel-Legendre transforms of G and

H:

$$G_x^*(kD) = \max_U \left\{ \sum_y \int \frac{k_x D^y(k)}{n_x} U_{xy}(k) - \sum_x G_x(U)\, dk \right\}$$

$$H_y^*(D) = \max_V m_y \left\{ \int \sum_y \frac{D^y(k)}{m_y} V_y(k) - \sum_y H_y(V)\, dk \right\}$$

For which Euler-Lagrange equations are

$$\frac{k_x D^y(k)}{n_x} = \frac{\exp(u_{xy}(k))}{\sum_y \int \exp(u_{xy}(k))\, dk} \text{ and } \frac{D^y(k)}{m_y} = \frac{\exp(v_y(k))}{\int \exp(v_y(k))\, dk}$$

Rearranging yields both (14) and (15). $\square$