22
WORKING
PAPERS 2023

WHAT LIES BEHIND RETURNS TO SCHOOLING: THE ROLE OF LABOR MARKET SORTING AND WORKER HETEROGENEITY

Pedro Portugal | Hugo Reis
Paulo Guimarães | Ana Rute Cardoso

BANCO DE PORTUGAL
EUROSYSTEM

22
WORKING
PAPERS 2023

WHAT LIES BEHIND RETURNS TO SCHOOLING: THE ROLE OF LABOR MARKET SORTING AND WORKER HETEROGENEITY

Pedro Portugal | Hugo Reis
Paulo Guimarães | Ana Rute Cardoso

DECEMBER 2023
The analyses, opinions and findings of these papers represent the views of the authors, they are not necessarily those of the Banco de Portugal or the Eurosystem.

Please address correspondence to
Banco de Portugal
Rua do Comércio 148, 1100-150 Lisboa, Portugal
Tel.: +351 213 130 000, email: info@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisboa, 2023 • www.bportugal.pt

Working Papers | Lisboa 2023 • Banco de Portugal Rua do Comércio 148 | 1100-150 Lisboa • www.bportugal.pt •
Edition Banco de Portugal • ISBN (online) 978-989-678-882-7 • ISSN (online) 2182-0422

What lies behind returns to schooling: the role of labor market sorting and worker heterogeneity

Pedro Portugal
Banco de Portugal
Nova SBE

Hugo Reis
Banco de Portugal
Católica Lisbon School of Business & Economics

Paulo Guimarães
Banco de Portugal
Universidade do Porto

Ana Rute Cardoso
Universidade de Lisboa
Instituto de Ciências Sociais

December 2023

Abstract
Do more educated workers earn higher wages partly because they have access to high-paying firms and occupations? We rely on linked employer-employee data on Portugal to combine the estimation of AKM models with the decomposition of the returns to schooling. We exploit exogenous variation in education driven by changes in compulsory education. We show that education provides access to better-paying workplaces and occupations: 30% of the overall return to education operates through the workplace channel and 12% through the occupation channel. The remainder is associated exclusively with the individual. Match quality plays a modest role in the returns to education.

JEL: I26; J31; J24

Keywords: wage distribution; returns to education; linked employer-employee data; high-dimensional fixed effects; workplace; occupation.

Acknowledgements: We would like to thank David Card, Till Von Wachter, and Juan Jimeno for their helpful comments, and the comments by participants in the Bank of Spain seminar, Bank of Portugal and Bank of Spain Workshop on Labor Economics, and in several seminars and conferences. The authors acknowledge the financial support of Fundação para a Ciência e Tecnologia through the project number PTDC/EGE-ECO/7493/2020. Pedro Portugal acknowledges financial support of Fundação para a Ciência e a Tecnologia (UIDB/00124/2020, UIDP/00124/2020 and Social Sciences DataLab - PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016). Cardoso acknowledges financial support of Fundação para a Ciência e a Tecnologia, contract 2020.03281.CEECIND/CP1615/CT0002. A significant part of this paper previously circulated in the Working Paper "The Returns to Schooling Unveiled", which was further expanded and subsequently split into two different articles. We thank Lucena Vieira for very competent research assistance. The analyses, opinions and conclusions expressed herein are the sole responsibility of the authors and do not necessarily reflect the opinions of Banco de Portugal or the Eurosystem.

1. Introduction

Two separate strands of literature have dealt with the returns to education, and wage heterogeneity across employers. We follow on these lines of research to explore the labor market returns to education adopting a comprehensive approach, which takes into account *who* the worker is (worker unobserved ability), *what* he does (the occupation), and, also crucially, *for whom* (the employer). The main aim of the analysis is therefore to pinpoint the role of firm and occupation heterogeneity in shaping the returns to schooling. Indeed, firms' pay standards are heterogeneous and workers with different levels of education are not randomly allocated to firms. To the extent that education can grant a "passport" to better paying firms, part of the overall return on education would operate through a firm channel. A similar argument could be built over occupations. We progress from Card *et al.* (2018), Card *et al.* (2013), Carneiro *et al.* (2012), and Card *et al.* (2016) to quantify the impact of sorting of workers across firms and occupations on the returns to education.

We rely on longitudinal data on the population of firms and workers in the Portuguese economy. This enables us to observe the entire distribution of characteristics and outcomes of individuals, together with the attributes of their firms. Our linked employer-employee dataset is valuable for additional reasons. First, it reports the schooling of the worker. The combination of full coverage of the economy with data on schooling had never before been used for research on the returns to education. Secondly, we have accurate information on hours worked, as well as a control variable on whether the worker's earnings refer to full schedule and full earnings during the month. Therefore, we can undertake an analysis of hourly wages. Finally, our earnings data are not subject to any type of top-coding.

We start with the estimation of a base wage regression that we augment to include sets of high-dimensional fixed effects, in the spirit of the seminal work of Abowd *et al.* (1999) (hereinafter AKM). We account for firm and occupation, as well as worker time-invariant heterogeneity. We do so in an OLS setting and, more relevantly, using instruments that exploit exogenous variation in education driven by changes in mandatory schooling in Portugal. We then adapt Gelbach (2016) unambiguous conditional decomposition of the impact of various omitted covariates on an estimated coefficient, to quantify how much of the return to education operates through a firm and an occupation channel, as opposed to a worker individual channel. The advantage of this procedure over the most often used alternative of adding covariates in sequence, is that Gelbach's decomposition yields unambiguous results and the order of inclusion of the covariates does not influence their contribution. The exercise undertaken can be interpreted very intuitively taking the example of the firm channel —it brings to light differences in firm wage effects across schooling levels. In other words, it quantifies the relevance of worker sorting across firms in shaping the returns to education. Engbom and Moser (2017) undertook a related, though much more limited exercise, as they estimated the returns to a bachelor, a master, and a PhD degree, comparing the results with and without firm fixed effects, in a wage regression that did not account for worker unobserved heterogeneity. We document that education provides access to better-paying firms and occupations: 30 percent of the overall return to education operates through the firm channel and 12 percent operates through the occupation channel, while the remainder is associated exclusively with the individual.

Next, we address an unsettled debate on the impact of worker-firm match effects as a determinant of wages. On the one hand, early work by Topel and Ward (1992) interpreted job mobility as a search for improved worker-firm match quality that would drive wage progression. Goldschmidt and Schmieder (2017) find that in Germany, the firm wage effects for workers in cleaning, security, and logistics services are lower than for other workers, mainly through a process of outsourcing. Eeckhout (2018) provides an overview of theoretical models of sorting and its implications. On the other hand, the work by Card *et al.* (2018) and Macis and Schivardi (2016) indicates that firms apply a consistent pay standard to all of its workforce —hence good firms are sought after by workers of all types, from the low-skilled to the high-skilled. As a result, the match effect brings little additional information over a model with worker and firm wage effects separately. We consider specifically the role of match effects in the returns to education. We find that match quality plays a negligible role in shaping the returns to education.

Section 2 provides an overview of the literature on wage heterogeneity across employers. Section 3 describes the institutional setting in the Portuguese labor market, followed by the data section 4. Sections 5 and 6 present the methodology and results on the impact of worker sorting across firms, and occupations structuring the returns to education. The role of the match quality is discussed in Section 7. Section 8 concludes.

2. Returns to Education: Current Evidence on the Role of the Employer

Evidence that firms may find it profitable to deviate from a market-wide wage standard was reported as early as the 1950s. Case studies by Lester (1952) and Reynolds (1951) have shown that employers' pay standards vary widely, even within narrowly defined regions and industries. Later, Groshen (1991a) documented a large contribution of the employer to intra-industry wage differentials. Machin and Manning (2004) corroborated the idea that wages are far from competitive, as they documented high wage dispersion across firms within a narrowly defined occupation and geographic area, despite the operation of a large number of firms delivering a homogeneous good. AKM started a very prolific line of literature that explores large longitudinal linked employer-employee data to quantify firm effects on wages. This strand now includes Gruetter and Lalive (2009), Eeckhout and Kircher (2011), Card *et al.* (2013), Torres *et al.* (2018), and Lopes de Melo (2018), among several others. Card *et al.* (2018) summarize the literature on the role of the firm in wage dispersion.[1] We are now equipped with several theoretical explanations about why firms may find it profitable to deviate from a market-wide wage standard: efficiency wages, implicit contracts, rent-sharing, principal-agent models, and the frictions contemplated in search and matching models (Addario *et al.*, 2022).

This empirical literature has remained silent on the impact of employer policies on the returns to schooling specifically (see the overviews of the state of the art research on human capital by (Deming, 2022; Card, 1999, 2001; Blundell *et al.*, 2005; Belzil, 2007)). Engbom and Moser (2017) attempted to quantify the impact of firms on the returns to education, but their dataset is restricted to higher education degrees and their analysis diverges from AKM because it does not account for worker effects. Similarly to the firm, the role of the job has also been neglected when studying the returns to education, despite early concern about the impact that the introduction of controls for broad occupation might have on the estimates.

3. Institutional Setting on Wages

A national minimum wage is enforced in Portugal, defined as a monthly rate for full-time work. Currently, sub-minimum wages apply only to physically disabled workers and trainees, after the abolition in 1999 of all reductions based on age.

Collective bargaining plays a central role in the Portuguese labor market, as in several other continental European economies. Massive collective agreements, often covering an industry, are common in the economy. Firm level collective bargaining traditionally covers a low share of the workforce, less than 10%. Extension mechanisms are common, either by mandatory government regulation or on a voluntary basis, as employers automatically apply the contents of collective agreements to their non-unionized workforce.

It should be noted that, despite the relevance of collective bargaining, firms have always enjoyed some degree of freedom in wage setting. Wage cushion (the difference between the actual wage level and the bargained wage level) promotes an alignment of wages with industry- and firm-level conditions, as documented in detail by Cardoso and Portugal (2005), Card and Cardoso (2022), and Addison *et al.* (2022). It follows from such an institutional setting that it is of key interest to quantify the impact of the firm when estimating the returns to education.

1. The impact of the employer determining the gender pay gap has received the attention of Groshen (1991b), Blau (1977), Meng and Meurs (2004), Card *et al.* (2016), and Cardoso *et al.* (2016).

4. Data Source and Concepts Used

Quadros de Pessoal (QP) is an unusually rich and comprehensive linked employer-employee data set, gathered annually by the Ministry of Employment. It covers all establishments having at least one wage earner. The wage information is collected with reference to the month of October. Civil servants, self-employed, and household employees are not covered; the share of wage-earners in agriculture is low and therefore the coverage of this sector is low. Instead, for manufacturing and the services private sector of the economy, the survey covers virtually the entire population of workers and firms.

The following variables are reported on each worker: gender, age, schooling, occupation, date of hire into the firm, monthly earnings, hours of work, the collective bargaining agreement, and the worker's job title. The education variable is defined as the number of years required to achieve the highest schooling degree. This variable is time-varying but only a small proportion of workers actually change their schooling degree once they enter the labor market. The share of workers for whom schooling varies over time is 14.7 percent, corresponding to 2.8 percent of the observations. Information on the employer includes the industry and location. In the current exercise we use information stretching from 1995 to 2021.[2]

The information on wages and their components is unusually trustworthy. The main reason why measurement error is attenuated in our dataset is because this administrative data was created to make sure that employers would comply with the wage rates negotiated in collective bargaining (Cardoso *et al.* 2016, p. 508).

We have restricted the analysis to workers aged 16 to 64, reported working full-time in the non-agricultural sectors, with at least 120 monthly hours of work, whose base wage does not fall below the national minimum wage, with non-missing schooling, and reported job duration between 0 and 600 months.

To separately identify firm and worker wage effects, the analysis must be restricted to the set of firms that are connected by worker mobility (see the discussion in Abowd *et al.* (2002)). We therefore limit our analysis to the largest connected set of observations defined as connected for three fixed effects (worker, firm, and occupation). The final dataset under analysis comprises 47.6 million observations on 6.4 million workers, 704 thousand firms, and 10596 occupations/year.[3]

Hourly wages are computed as the actual overall monthly earnings (including base wage, tenure-related and other regularly paid components) over the number of normal hours of work. Wages were deflated using the consumer price index (base 2021), but this correction is inconsequential since we will always use year dummies.

2. Notice that no worker data is available for 2001.

3. We use occupation-year fixed effects to circumvent changes in the occupation classification over the sample period.

Table A.1.1 in the Appendix presents the descriptive statistics for the variables used in the estimation.

5. Sorting of Workers across Firms and the Returns to Education

5.1. Wage regressions with worker and firm effects

We begin by specifying a standard Mincerian wage equation with two added fixed effects for worker and firm. This popular specification was introduced by Abowd *et al.* (1999) (hereafter, AKM) and requires the use of linked employer-employee data. By controlling for firm and worker fixed effects we are able to control for unobservables at the firm and worker level that capture a substantial amount of wage variation while mitigating potential endogeneity problems. More specifically, we consider an equation of the type,

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha_i + \theta_{\mathbf{F}(i,t)} + \mu_t + \varepsilon_{it} \quad , \tag{1}$$

where y_{it} is the logarithm of the hourly wage for each worker i $(i = 1, ..., N)$ at year t $(t = 1, ..., T)$; $\mathbf{x}_{it}$ is a vector of observed time-varying characteristics of workers; α_i is a time-invariant worker fixed effect; $\theta_{\mathbf{F}(i,t)}$ is a firm specific time-invariant fixed effect; μ_t are time fixed effects; and ε_{it} is the disturbance term of the regression.[4] The vector of explanatory variables, $\mathbf{x}_{it}$, comprises a quadratic on age of the worker, a quadratic on tenure and years of schooling. Gender is time invariant and is only explicitly accounted for in specifications that omit the worker fixed effect. The estimated regression coefficients are unbiased under the assumption of strict exogeneity, $E(\varepsilon_{it}|\mathbf{x}_{it}, \alpha_i, \theta_{\mathbf{F}(i,t)}, \mu_t) = 0$.

Estimation of equation (1) by ordinary least squares (OLS) is complicated by the fact that it includes two high-dimensional fixed effects. As discussed in AKM, the large dimension of the design matrices for the fixed effects makes impractical the application of the conventional OLS formula. Fortunately, models of this type can be estimated using, for example, the algorithm proposed by Guimarães and Portugal (2010).[5]

Basically, the algorithm consists of an iterative procedure that alternates between the estimation of the fixed effects (taking as given the last estimates of the $\boldsymbol{\beta}$) and estimation of $\boldsymbol{\beta}$ (taking as given the last estimates of the fixed effects). This algorithm has the advantage of converging (albeit at a slow rate) to the true OLS solution. There is, however, an additional complication that arises in models with more than one high-dimensional fixed effect. The likely existence of

4. The parentheses in the subscripts of the fixed effects coefficients are used to emphasize that the ultimate source of variation stems from the worker/time combination.

5. The user-written package reghdfe written by Sérgio Correia and available on the Statistical Software Components (SSC) Boston Archive implements an improved version of this algorithm. This package allows for estimation of models with multiple high-dimensional fixed effects.

perfect multicollinearity between parameters associated with the fixed effects may introduce problems of identification. This may not be an issue if interest centers on the $\boldsymbol{\beta}$ coefficients but in our case we also want to implement secondary analysis of the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. Interpretation of the estimates of the fixed effects is only meaningful if the differences between coefficients (within each fixed effect) are estimable. To guarantee identification, and thus ensure comparability of the parameter estimates, we restrict our analysis to the largest subset of data where all the fixed effects are connected (the largest mobility group on Abowd *et al.* (2002) parlance).[6]

5.2. Gelbach's decomposition

To understand the contribution that the allocation of workers to firms has to the observed education pay differential we make use of Gelbach (2016) decomposition method. His approach is based on the OLS formula for omitted variable bias and allows for a decomposition that unambiguously quantifies the portion of the variation attributed to each variable of interest. Gelbach's decomposition is easier to present if we resort to matrix notation. Consider a conventional Mincerian equation that includes the observable characteristics as well as time effects. For convenience, we collect the observations for all variables but worker schooling, into the matrix $\mathbf{Z}$. Our variable of interest, schooling, is introduced separately and represented by the variable $\mathbf{S}$. Thus, we have

$$\mathbf{Y} = \delta_0 \mathbf{S} + \mathbf{Z}\boldsymbol{\gamma_0} + \varepsilon \quad . \tag{2}$$

By the Frisch-Waugh-Lovell theorem we know that the same OLS estimate of δ_0 may be obtained by running a simple regression of $\mathbf{Y}$ on $\mathbf{S}$ after partialing out the effect of $\mathbf{Z}$ from both variables. More specifically,

$$\widehat{\delta_0} = \left(\mathbf{S}'\mathbf{M_Z}\mathbf{S}\right)^{-1}\mathbf{S}'\mathbf{M_Z}\mathbf{Y} = \mathbf{H_Z}\mathbf{Y} \quad , \tag{3}$$

where $\mathbf{M_Z} \equiv \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ is the well-known symmetric and idempotent residual-maker matrix. Here $\widehat{\delta_0}$ is the conventional OLS estimator used to produce estimates for the returns to education. To show how Gelbach's decomposition can be used to tease out the contribution of the worker and firm fixed effects on the returns to education, consider now the full regression in (1) which is the same as (2) with two added sets of fixed effects: worker ($\boldsymbol{\alpha}$) and firm ($\boldsymbol{\theta}$). This regression, written in terms of its fitted OLS expression, is:

$$\mathbf{Y} = \widehat{\delta_1}\mathbf{S} + \mathbf{Z}\widehat{\boldsymbol{\gamma}} + \mathbf{D}\widehat{\boldsymbol{\alpha}} + \mathbf{L}\widehat{\boldsymbol{\theta}} + \mathbf{e} \quad . \tag{4}$$

6. We used the algorithm described in Abowd *et al.* (2002) to identify the mobility groups. The largest mobility group accounted for over 98% of our original data set thus rendering negligible possible concerns about sample selection bias.

where $\mathbf{e}$ are the regression residuals. Note that $\mathbf{D}\widehat{\boldsymbol{\alpha}}$ and $\mathbf{L}\widehat{\boldsymbol{\theta}}$ are column vectors containing the least-squares estimates for the worker and firm fixed effects in a regression that also controls for $\mathbf{S}$ and $\mathbf{Z}$. To obtain a decomposition of $\widehat{\delta_0}$ we multiply both terms of equation (4) by $\mathbf{H_Z}$. In other words, we regress each element of the above equation on education while controlling for the remaining observable variables ($\mathbf{Z}$). On the left-hand side we obtain $\widehat{\delta_0}$ directly and, given that $\mathbf{H_Z Z}\widehat{\gamma} = \mathbf{0}$ and $\mathbf{H_Z e} = \mathbf{0}$, the right-hand side is simply:

$$\widehat{\delta_0} = \widehat{\delta_1} + \mathbf{H_Z D}\widehat{\boldsymbol{\alpha}} + \mathbf{H_Z L}\widehat{\boldsymbol{\theta}} = \widehat{\delta_1} + \widehat{\delta}_{\boldsymbol{\alpha}} + \widehat{\delta}_{\boldsymbol{\theta}} \quad . \tag{5}$$

This means that the conventional return on education, $\widehat{\delta_0}$, can be decomposed into three terms that reflect the impact of the worker and firm channel. If, conditional on all $\mathbf{Z}$ covariates, workers were randomly allocated to firms, then the estimate for $\widehat{\delta}_{\boldsymbol{\theta}}$ would be zero. In this case the distribution of schooling levels within each firm cell would replicate the distribution of schooling levels in the economy, such that the matching of schooling levels to firm with different pay standard would not be a source of returns to education. On the other hand, a positive value for $\widehat{\delta}_{\boldsymbol{\theta}}$ would be a clear indication that better-educated workers are sorted to higher-paying workplaces. From the equation above we see that the estimate of $\widehat{\delta}_{\boldsymbol{\theta}}$ may be interpreted as the log point reduction/increase that occurs in the returns to schooling due to the allocation of workers to firms. Gelbach's decomposition holds as well with instrumental variables where we need only to replace the endogenous regressor(s) by their instrumented values.[7] In our case, since instrumented education is time-invariant, $\widehat{\delta_1}$ in equation (5) will be absorbed by $\widehat{\delta_\alpha}$.

5.3. Empirical results on benchmark regression

We start by estimating a conventional OLS human capital wage function including as covariates a quadratic term on age of the worker, a quadratic term on tenure, the worker gender and schooling, together with year fixed effects. Table 1 reports the results of the OLS specification in Column (1).

As expected, wages increase with age and tenure at a decreasing rate, reaching the maximum at 61 and 49 years, respectively. The female gender wage gap in Portugal over this period is estimated to be 24.0 percent (-27.4 log points). In Portugal, each additional year of education yields, on average, an 8.9 percent (8.5 log points) labor market return. This return is in line with international evidence, even though it places Portugal among the countries with relatively high returns to schooling (Harmon *et al.*, 2003; Card, 1999); the cross-country survey of estimates

7. According to Gelbach (2016), it can be shown that under exact identification, the decomposition holds when two stage least square are employed. "When β_1 is overidentified the equivalence does not hold in fixed samples but it does hold asymptotically, provided all instruments are valid." (Gelbach 2016, p. 526)

	OLS		IV	
	Base	Full	Base	Full
	(1)	(2)	(3)	(4)
Age	0.0367	-	0.0358	-
	(0.0001)	-	(0.0001)	-
Age Squared	-0.0003	-0.0003	-0.0003	-0.0003
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Tenure	0.0196	0.0102	0.0196	0.0100
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Tenure squared	-0.0002	-0.0002	-0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gender(Female=1)	-0.2741	-	-0.2792	-
	(0.0004)	-	(0.0005)	-
Schooling	0.0854	0.0070	0.0934	-
	(0.0001)	(0.0001)	(0.0006)	-
Time effects	✓	✓	✓	✓
Worker effects		✓		✓
Firm effects		✓		✓
N	47,569,720	47,569,720	47,569,720	47,569,720
R^2	0.4272	0.8856	0.4247	0.8855

Table 1. Wage Equations, OLS and IV

Notes: The dependent variable is the logarithm of real hourly wages. Column (1) reports the OLS results of the benchmark specification including as covariates age, age squared, tenure, tenure squared, gender, worker schooling, and year fixed effects. Column (2) shows the full specification, including worker and firm fixed effects. In this full specification age and gender are absorbed by the worker-fixed effects and time effects. Column (3) reports the IV results of the benchmark specification, using the changes in the compulsory schooling years as an instrument to the worker years of schooling. Column (4) shows the full specification for the IV approach, including worker, firm, and year fixed effects. In this case, on top of age and gender, schooling is also absorbed by the worker fixed effects due to the invariant nature of the instrument at the individual level. Standard errors are clustered at the worker level.

by Ashenfelter *et al.* (1999); Trostel *et al.* (2002); and Montenegro and Patrinos (2014)).

Nevertheless, that figure may fail to convey the causal effect of education on wages, as widely acknowledged in the literature (see the summary and discussion in Card (1999)). In general, more productive individuals may find it more advantageous to invest in education, which will confound the estimation of the causal returns to education and lead to a positive bias in an OLS setting (ability bias). If schooling is subject to measurement error, the estimation will be biased in the opposite direction (attenuation bias). Moreover, schooling is the result of optimizing behavior on the part of individuals, who are likely to choose the kinds of study for which they are most motivated and competent or where they expect to obtain the highest benefit. Thus, in the presence of individual heterogeneity, the OLS return to education can be over or underestimated (selection bias). In our specific setting, the concerns may go further. The sorting of better educated workers into better employers, may be due to education per se, or due to unobserved factors. In particular, individuals with higher capability to find a good employer may systematically also find it more advantageous to invest in education, which will confound the estimation of the impact of the employer channel on wages that we aim for. Tackling this problem requires exogenous variation in schooling that can be exploited in the empirical model. Such source of variation would enable the quantification of the impact of education, purged of individual unobserved capabilities that may jointly affect the schooling attainment, the wage, and the quality of the individual's employer.

Changes in compulsory schooling laws yield exogenous variation in schooling that has been widely used in the literature (Acemoglu and Angrist (2000)). The legal setting on compulsory education evolved over the birth cohorts in our analysis, as described in Table 2. In 1911, 3 years of elementary education became compulsory for children of both genders aged 7 to 14. Throughout the 1940s and the 1950s, an investment program to build elementary schools was implemented and an effort was made to speed up the training of teachers. In 1954, *Plano de Educação Popular*, a comprehensive government program aiming at increasing alphabetization of children (and adults) was enacted. Within its scope, a system of sanctions was created to penalize those not obeying the compulsory education laws. Two years later, in 1956, compulsory schooling for boys increased to 4 years. This was extended to girls in 1960. Compulsory education was then extended to 6 years in 1964. In 1986, it was raised to 9 years, and more recently, in 2009, the minimum compulsory education level was set at 12 years.

The specification of the first stage equation uses as instruments the dummy variables, which makes the correspondence between the compulsory schooling rules with the birth date of each individual. The estimation results of the first stage are reported in Table A.1.2.

Column (3) in Table 1 reports the IV results. The regression coefficient estimates are higher once we instrument for education relying on changes in compulsory schooling laws, in line with the literature results. The IV estimate of the returns to education is now 9.8 percent (9.34 log points), nearly one percentage point higher than the one obtained by OLS. This seems to suggest that the marginal returns to education for those affected by the changes in the compulsory schooling

Year	Document	Years of compulsory schooling	First cohort affected	School entry age
1911	DL from March 29	3 years	Boys and girls born after 1904 (to be enrolled in the 1st grade in 1911)	7
1956	DL 40964	4 years	Boys born in or after 1950 (to be enrolled in the 1st grade in 1957)	7
1960	DL 42994	4 years	Girls born in or after 1953 (to be enrolled in the 1st grade in 1960)	7
1964	DL 45810	6 years	Boys and girls born in or after 1957 (to be enrolled in the 1st grade in 1964)	7
1986	Law 46/86	9 years	Boys and girls born in or after 1980 (to be enrolled in the 1st grade in 1986)	6
2009	Law 85/2009	12 years	Boys and girls born in or after 1998 (enrolled up to the 7th grade in 2009)	6

Table 2. Key legislation on compulsory schooling

Source: *Diário da República*.

laws are higher when compared to the average marginal returns to education in the population as a whole.

Next, we report the full specification of our initial empirical model, which includes worker and firm fixed effects (columns (2) and (4) in table 1). Figures 1 and 2 depict the raw wages in the economy, as well as an overview of the different sets of fixed effects, separately for three educational levels: basic education, secondary education, and college education.[8] Raw wages for the lowest education level are, as expected, lower, but they are as well more concentrated than for any other educational group. This relatively low dispersion of wages could reflect the operation of collective bargaining, setting binding wage floors for low-skilled workers, and in particular, the role of mandatory minimum wages. College education, instead, yields the most heterogeneous returns in the economy.

8. According to the Portuguese system, basic education corresponds to 9 years of schooling and secondary stands for 12 years of schooling. The college group corresponds to those individuals that completed a degree after secondary education, namely an undergraduate, a master or a PhD program.

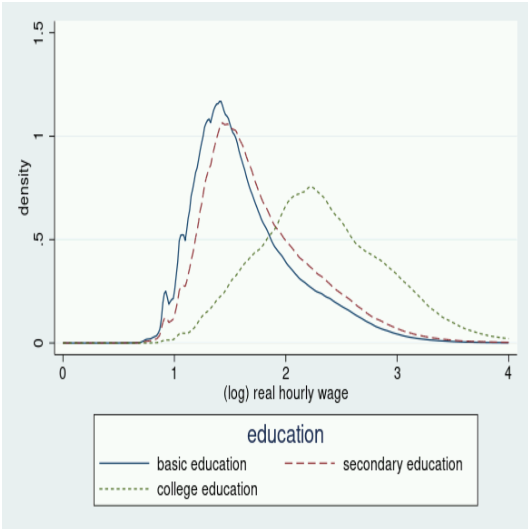

Figure 1: Distribution of (log) wages, separately by education level

Notes: Reports kernel densities of log hourly wages in the economy separately for three educational levels: basic education, secondary education, and college education.

Figure 2 illustrates the heterogeneity of wage policies across firms through the firm fixed effects. The distribution reveals the existence of a wide range of pay standards across firms and the presence of mass points that correspond to large firms in the economy. In addition, there is some evidence that more educated workers have better access to higher-paying firms.

The worker-fixed effects represent the permanent worker heterogeneity, both observed and unobserved. A high worker fixed effect or high-wage worker is an individual with total compensation higher than expected after controlling for observable time-varying worker, and for firm permanent heterogeneity. It is clear that higher worker fixed effects tend to be associated with higher levels of schooling. Furthermore, the dispersion of worker abilities is considerably larger among college graduates than among the other schooling levels.[9]

We next provide evidence of the importance of mandatory schooling, which corresponds to our instrumental variable, driving the distribution of the worker and firm fixed effects. Figure 3 shows that cohorts with higher levels of mandatory schooling (four, six, and nine years of compulsory schooling) are associated with higher fixed effects, suggesting that mandatory schooling may serve as a useful instrument.

9. Some key statistical moments of the wage distribution, including variance decomposition, correlations, and fixed effects heterogeneity, are provided in Table A.1.3 in the Appendix.

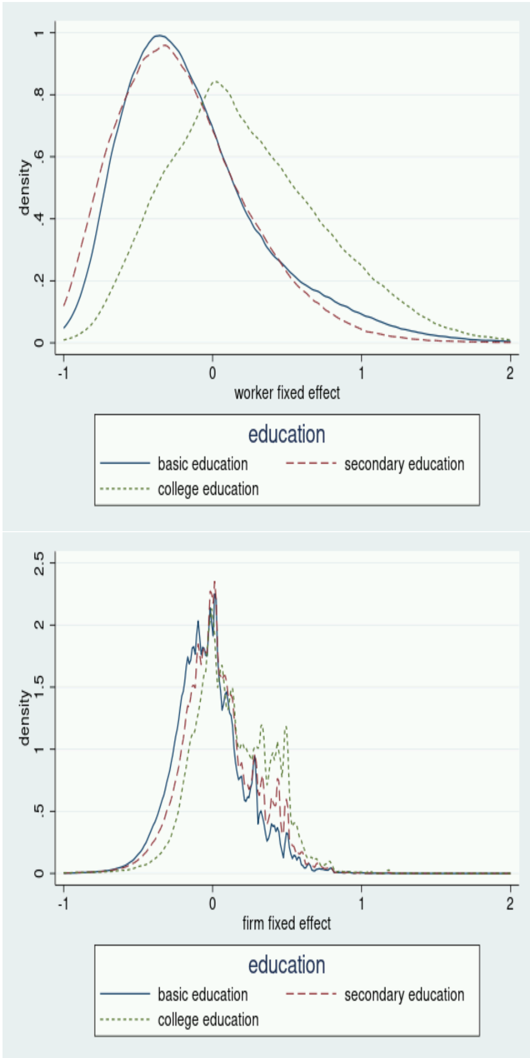

Figure 2: Distribution of firm and worker fixed effects, separately by education level

Notes: The left-side panel plots the kernel densities for the worker fixed effects and the right-side panel plots the firm fixed effects, separately for the three educational levels. These figures follow from the estimation reported in Table 1 Column (2).

In the following section, we quantify precisely the impact of the distinct channels determining the returns to education.

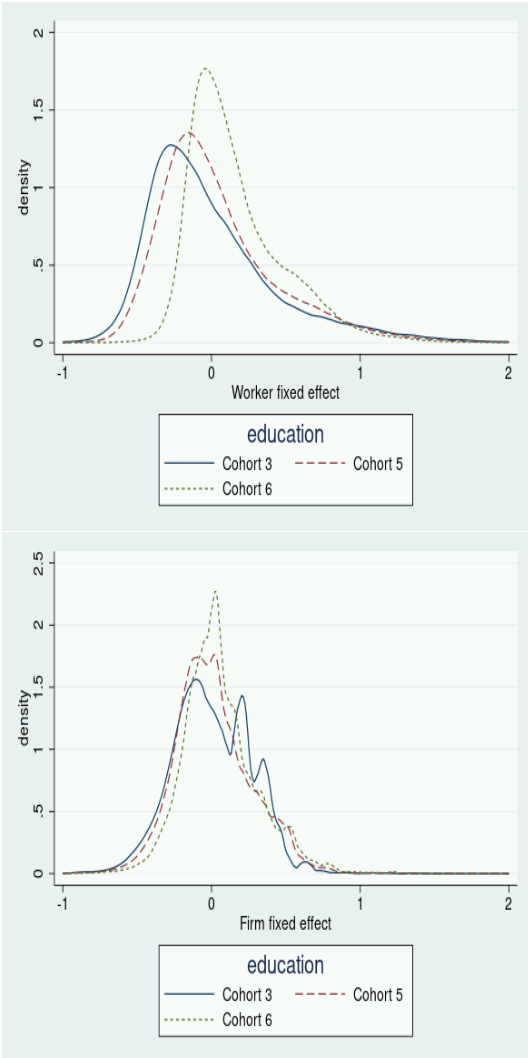

Figure 3: Distribution of Worker and Firm fixed effects, separately by cohorts with different mandatory schooling laws

Notes: Reports kernel densities of worker and firm fixed effects for three cohorts: Cohort 3 - Males born between 1950 and 1956 (4 years of compulsory schooling), Cohort 5 - Males and Females born between 1957 and 1979 (6 years of compulsory schooling), and Cohort 6 - Males and Females born between 1980 and 1997 (9 years of compulsory schooling). To guarantee comparability we have fixed the age of the individuals to be 40.

5.4. Decomposing returns to education

We now quantify the relevance of the different channels shaping the returns to education previously estimated. The first is the employer channel, which operates to the extent that education provides entry to workplaces with more generous pay standards. This mechanism operates as long as the allocation of workers with different schooling levels is not orthogonal to the workplaces' pay standard. It thus reflects the existence of systematic sorting of educational levels across firms.

The remaining channel, after accounting for firm heterogeneity in pay standards, would be the individual component of the returns to education. Such component encompasses both a "pure" return on the worker's education and a return on other individual attributes, whether observed or unobserved.

Panel A - Gelbach Decomposition of the Return to Education - OLS

Benchmark	Full	Decomposition into:	
Regression	Specification	Worker FE	Firm FE
(1)	(2)	(3)	(4)
0.0854	0.0070	0.0564	0.0220
(0.00006)	(0.00006)	(0.00005)	(0.00088)

Panel B - Gelbach Decomposition of the Return to Education - IV

Benchmark	Full	Decomposition into:	
Regression	Specification	Worker FE	Firm FE
(1)	(2)	(3)	(4)
0.0934	0.0000	0.0652	0.0282
(0.00059)	-	(0.00047)	(0.00180)

Table 3. Conditional Decomposition of the Return to Education - OLS and IV

Notes: Panel A shows Gelbach's conditional decomposition of the return to education based on the OLS regression. Column (1) reports the coefficient of the benchmark result on return to education. Column (2) reports the estimate for the coefficient on education of the full specification after including worker and firm fixed effects. The results of the decomposition are reported in columns 3 and 4. Adding up the results of Columns (2), (3), and (4) we obtain the benchmark coefficient in Column (1). Panel B reports equivalent results for the IV case. Since instrumented schooling is time-invariant it is absorbed by the worker-fixed effect. Standard errors are clustered at the corresponding group level (worker or firm).

Table 3 reports the OLS and the IV results from the Gelbach decomposition, starting with a benchmark regression where we exploit exogenous variation in education. Focusing on the IV results, column (1) of panel B shows the coefficient of the benchmark IV estimate on the returns to education. Column (2) reports the coefficient of the full specification that includes worker and firm fixed effects, which is zero by construction due to the invariant nature of the instrument at the individual level. The results of the decomposition are reported in Columns (3) and

(4) of panel B, which, by construction, add up to the coefficient of the benchmark specification.

We find that 70 percent of the return to education is an individual component purged of sorting into firms (6.52 log points out of the 9.34 overall return on education). In other words, this decomposition shows that the economy's return to education would fall by 2.82 log points if workers of different schooling levels were randomly distributed across firms. Therefore, almost one third of the returns to education operates via the allocation of workers to firms —"firm channel".[10] The reader may be concerned that the time invariance of the firm fixed effects is too restrictive. However, firm fixed effects are highly persistent over time. Considering a specification with time-varying firm fixed effects does not change the contribution of the worker and firm channels to the returns to education (6.53 log points and 2.84 log points).

We also addressed a possible concern with the incidental parameter bias engendered by the presence of high-dimensional fixed effects (Bonhomme *et al.*, 2023; Carneiro *et al.*, 2023). In our case, under the classical assumptions, we do not expect that the inclusion of high dimensional fixed effects engenders any bias in the estimate of the regression coefficients. In line with our expectations, the results of the split-panel Jackknife bias correction procedure (Dhaene and Jochmans, 2015) suggest that, the incidental parameter problem is not materially important, neither in the OLS nor IV estimates.[11] [12]

6. Sorting of Workers across Firms and Occupations and the Returns to Education

6.1. Wage regression with worker, firm, and occupation effects

To capture the potential role of sorting across occupations we add to equation (1) an additional fixed effect that captures wage heterogeneity due to occupations. Thus, our Mincerian wage equation includes now three high-dimensional fixed effects: worker, firm, and occupation. More specifically, we consider an equation of

10. Table 3, panel A, reports a comparable decomposition exercise, using the OLS regression. We find similar results under the OLS. In relative terms, the worker component of the return to education increases, as we compound it with the worker unobserved ability. Consequently, the role of the firm channel decreases.

11. For simplicity, the bias correction estimates are available upon request.

12. It is well known that the incidental parameter problem tends to generate an upward bias of the variance of the estimated fixed effects. With multiple high dimensional fixed effects, sampling error in one dimension tends to be compensated with measurement error of opposite sign in other dimensions, inducing a negative correlation between the two. We use the variance/covariance bias correction methods proposed by Andrews *et al.* (2008) and Kline *et al.* (2020) but show that, in our data, the estimated biases are small. See Table A.1.4 in the Appendix.

the type,

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha_i + \theta_{\mathbf{F}(i,t)} + \lambda_{\mathbf{O}(i,t)} + \mu_t + \varepsilon_{it} \quad , \tag{6}$$

where $\lambda_{\mathbf{O}(i,t)}$ is an occupation specific time-invariant fixed effect and everything else remains as defined in equation (1). The inclusion of an additional high-dimensional fixed-effect does not present any complication for estimation because the algorithm of Guimarães and Portugal (2010) can easily be extended to more than two high-dimensional fixed effects.[13] The only practical consequence is the need to redefine the largest connected set. We use the algorithm of Weeks and Williams (1964) to identify a connected set. This algorithm can be applied when dealing with two or more sets of fixed effects and will produce the same result as the algorithm described in Abowd, Creecy, and Kramarz (2002) if applied to a model with two high-dimensional fixed effects.[14]

6.2. Empirical results accounting for worker, firm, and occupation unobserved heterogeneity

As described before, if workers endowed with better schooling levels are matched to better-paying firms, that will result in an education premium that we capture as firm channel, reflecting the existence of sorting of educational levels across firms. A strictly parallel reasoning would apply to occupations. If workers endowed with better schooling levels are matched to better-paying occupations, that will result in an education premium that we capture as the occupation channel, reflecting the existence of sorting of schooling levels across occupations. One would expect, of course, that the level of education plays a key instrumental role facilitating access to different occupations.[15]

In some sense a workplace can be seen as a collection of occupations. Different technologies and/or distinct human resources management strategies may result in combinations of high paying workplaces with high paying occupations. For example, it can be argued that technologically sophisticated firms often organize highly complex tasks. The empirical relevance of this “sophistication technology channel" should manifest itself via the association of the levels of education with the sign and magnitude of the assortative match between high paying firms and high paying occupations.

In this case, the remaining channel, after accounting for firm and occupation heterogeneity in pay standards, would be the individual component of the returns to education. As before, such component encompasses both a "pure" return on

13. As already noted the Stata package *reghdfe* will estimate linear regression models with multiple fixed effects.

14. The largest mobility group accounted for over 98% of our original data set, thus rendering negligible possible concerns about sample selection bias.

15. This relates to an old debate discussing whether one should control for occupation in a Mincerian regression when estimating the returns to education.

the worker's education and a return on other individual time-invariant attributes, whether observed or unobserved.

We now extend the full specification including worker, firm, and occupational fixed effects. Columns (2) and (4) in Table 4 present the OLS and IV results. Figure 4 provides the occupation fixed effect separately for the three educational levels, in line with what we have seen in Figure 2 for the worker and firm fixed effects. The occupation effects show a similar pattern to the one observed for firms, in particular, the existence of a wide range of pay standards across occupations.

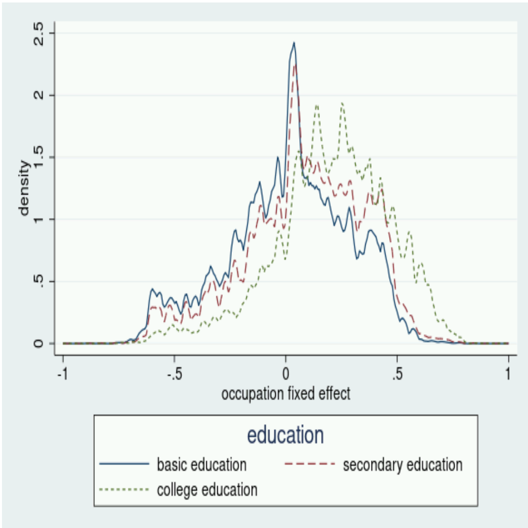

Figure 4: Distribution of occupation fixed effects, separately by education level

Notes: Plots the kernel densities for the occupation fixed effects, separately for the three educational levels. The figure follows from the estimation reported in Table 4 Column (2).

6.3. Decomposing returns to education

The aim of the current step of the analysis is to quantify how much of the education pay differential operates through the allocation of workers to firms and occupations. The decomposition presented in (5) for a full regression model that includes worker and firm effects can be easily adapted to the present case. If we add an occupation specific fixed effect (λ) to the equation in (4) then the decomposition of the return to education in the full equation becomes

$$\widehat{\delta_0} = \widehat{\delta_1} + \widehat{\delta}_{\boldsymbol{\alpha}} + \widehat{\delta}_{\boldsymbol{\theta}} + \widehat{\delta}_{\boldsymbol{\lambda}} \quad . \tag{7}$$

Equation (7) identifies the conditional contribution of firms and occupations to the returns to schooling, as well as the contribution of the worker-specific component. Table 5 reports the OLS (Panel A) and the IV (Panel B) results.

	OLS		IV	
	Base	Full	Base	Full
	(1)	(2)	(3)	(4)
Age	0.0367	-	0.0358	-
	(0.0001)	-	(0.0001)	-
Age Squared	-0.0003	-0.0003	-0.0003	-0.0003
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Tenure	0.0196	0.0097	0.0196	0.0096
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Tenure squared	-0.0002	-0.0002	-0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gender(Female=1)	-0.2741	-	-0.2792	-
	(0.0004)	-	(0.0005)	-
Schooling	0.0854	0.0053	0.0934	-
	(0.0001)	(0.0001)	(0.0006)	-
Time effects	✓		✓	
Worker effects		✓		✓
Firm effects		✓		✓
Occupation/time effects		✓		✓
N	47,569,720	47,569,720	47,569,720	47,569,720
R^2	0.4272	0.8900	0.4247	0.8899

Table 4. Wage Equations, OLS and IV (including occupation)

Notes: The dependent variable is the logarithm of real hourly wages. Column (1) reports the OLS results of the benchmark specification including as covariates age, age squared, tenure, tenure squared, gender, worker schooling, and year fixed effects. Column (2) shows the full specification, including worker, firm, and occupation/year fixed effects. In this full specification age and gender are absorbed by the worker fixed effects and time effects. Column (3) reports the IV results of the benchmark specification, using the changes in the compulsory schooling years as an instrument to the worker years of schooling. Column (4) shows the full specification for the IV approach, including worker, firm, and occupation/year fixed effects. In this case, on top of age and gender, schooling is also absorbed by the worker fixed effects due to the invariant nature of the instrument at the individual level. Standard errors are clustered at the worker level.

Column (1) shows the coefficient of the benchmark specification on returns to education. Column (2) provides the coefficient of the full specification including worker, firm, and occupation fixed effects. The results of the Gelbach decomposition are reported in columns (3) to (5).

Panel A - Gelbach Decomposition of the Return to Education - OLS

Benchmark	Full	Decomposition into:		
Regression	Specification	Worker FE	Firm FE	Occ FE
(1)	(2)	(3)	(4)	(5)
0.0854	0.0053	0.0491	0.0213	0.0097
(0.00006)	(0.00006)	(0.00005)	(0.00086)	(0.00026)

Panel B - Gelbach Decomposition of the Return to Education - IV

Benchmark	Full	Decomposition into:		
Regression	Specification	Worker FE	Firm FE	Occ FE
(1)	(2)	(3)	(4)	(5)
0.0934	0.0000	0.0541	0.0278	0.0115
(0.0006)	-	(0.0004)	(0.0016)	(0.0008)

Table 5. Conditional Decomposition of the Return to Education - OLS and IV (including occupation)

Notes: Panel A shows Gelbach's conditional decomposition of the return to education based on the OLS regression. Column (1) reports the coefficient of the benchmark result on return to education. Column (2) reports the estimate for the coefficient on education of the full specification after including worker, firm, and occupation/year fixed effects. The results of the decomposition are reported in columns 3 to 5. Adding up the results of Columns (2), (3), (4), and (5) we obtain the benchmark coefficient in Column (1). Panel B reports equivalent results for the IV case. Since instrumented schooling is time-invariant it is absorbed by the worker-fixed effect. Standard errors are clustered at the corresponding group level (worker, firm, or occupation).

Focusing on the IV regression, having estimated three sets of fixed effects —the firm, occupation, as well as the worker —we find, first of all, that almost 60% of the return to education is an individual component purged of sorting into firms and occupations (5.41 log points out of the 9.34 overall return to education). Secondly, the decomposition shows that the economy's return on education would decline by 2.78 log points if workers of different schooling levels were evenly distributed across firms. Therefore, thirty percent of the returns to education operates via the allocation of workers to firms —firm channel.

The allocation of workers to occupations has a smaller impact than the allocation to firms on the returns to education. Table 5 shows that the returns to schooling would decline by 1.15 log points if workers of different schooling levels were equally distributed across occupations, conditional on all other variables included in the full model. Therefore, around one eighth of the returns to education operates via the allocation of workers to occupations —occupation channel.

Sorting across occupations does not seem to play an important role explaining the returns to education. Following the approach in Gelbach (2016), it can be shown that although workers do sort into occupations by schooling (the results from a multivariate regression show that education plays a significant role allocating worker to occupations), and although occupations do pay different amounts (occupational wage differences are highly statistical significant), it turns out that most of the occupational sorting is weakly correlated with occupational wage premia.

The role of the firm's pay standards shaping wage differentials across education groups can be compared to its role shaping the gender pay gap. Cardoso *et al.* (2016) and Card *et al.* (2016) report a firm contribution to the gender pay gap around one fifth of the overall gap. We uncover that the role of the firm shaping the returns to education is larger than its role shaping the gender pay gap. To our knowlewdge, this is a novel fact that had so far deserved hardly any discussion in the literature (except the comment by Card *et al.* (2013) when dealing with Germany). Having come such a long way in recent decades, the literature on the returns to schooling had, nevertheless, not yet uncovered differences in firm wage effects across schooling levels.

7. Is the Match Quality Important for the Returns to Schooling?

7.1. Empirical results accounting for the match quality and occupation unobserved heterogeneity

In this section, we present a full specification including firm-worker match effects. More specifically, we consider an equation of the type,

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \lambda_{\mathbf{O}(i,t)} + \varphi_{\mathbf{i}\times\mathbf{F}} + \mu_t + \varepsilon_{it} \quad , \tag{8}$$

where $\varphi_{\mathbf{i}\times\mathbf{F}}$ is a fixed-effect that accounts for all firm-worker unique combinations. Adding this fixed-effect is equivalent to including three separate effects – a worker, a firm, and a firm-worker interaction fixed effect (a pure matching effect). Thus, the above model nests that presented in the previous section in the sense that it includes an additional firm-worker matching effect.

The regression results in Table 6 are motivated by the notion that better educated workers may potentially find better quality matches. Therefore, our full specification, includes time, occupation, and match fixed effects. The usefulness of this exercise is better understood when we proceed with the decomposition exercise in section 7.3.

	OLS		IV	
	Base	Full	Base	Full
	(1)	(2)	(3)	(4)
Age	0.0367	-	0.0358	-
	(0.0001)	-	(0.0001)	-
Age Squared	-0.0003	-0.0003	-0.0003	-0.0003
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Tenure	0.0196	0.0063	0.0196	0.0062
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Tenure squared	-0.0002	-0.0002	-0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gender(Female=1)	-0.2741	-	-0.2792	-
	(0.0004)	-	(0.0005)	-
Schooling	0.0854	0.0037	0.0934	-
	(0.0001)	(0.0001)	(0.0006)	-
Time effects	✓		✓	
Match effects		✓		✓
Occupation/time effects		✓		✓
N	47,569,720	47,569,720	47,569,720	47,569,720
R^2	0.4272	0.9251	0.4247	0.9251

Table 6. Wage Equations, OLS and IV (including occupation and match)

Notes: The dependent variable is the logarithm of real hourly wages. Column (1) reports the OLS results of the benchmark specification including as covariates age, age squared, tenure, tenure squared, gender, worker schooling, and year fixed effects. Column (2) shows the full specification, including match, and occupation/year fixed effects. In this full specification age and gender are absorbed by the worker-fixed effects and time effects. Column (3) reports the IV results of the benchmark specification, using the changes in the compulsory schooling years as an instrument to the worker years of schooling. Column (4) shows the full specification for the IV approach, including match, and occupation/year-fixed effects. In this case, on top of age and gender, schooling is also absorbed by the worker fixed effects due to the invariant nature of the instrument at the individual level. Standard errors are clustered at the worker level.

7.2. Gelbach decomposition

Since equation (8) only includes two fixed-effects the Gelbach's decomposition presented in (5) is functionally equivalent (with the necessary adaptations) to the present case. Thus,

$$\widehat{\delta}_0 = \widehat{\delta}_1 + \widehat{\delta}_{\boldsymbol{\lambda}} + \widehat{\delta}_{\boldsymbol{\varphi}} \quad , \tag{9}$$

where $\widehat{\delta}_{\boldsymbol{\varphi}}$ is obtained as $\mathbf{H_Z M}\widehat{\boldsymbol{\varphi}}$ and $\mathbf{M}$ is the design matrix for the firm-worker matching effect. Equation 9 shows that the conventional return on education, $\widehat{\delta}_0$, can be decomposed into three terms that reflect the impact of the worker-firm combination and the occupation channel. An estimate of zero for $\widehat{\delta}_{\boldsymbol{\varphi}}$ would be an indication that schooling levels were unrelated to the worker-firm effect while a positive value would provide evidence that the sorting of more educated workers to higher firm-worker matches was a source of returns to education.

Is it possible to go further and decompose $\widehat{\delta}_{\boldsymbol{\varphi}}$ on the contribution due to the worker, the firm, and a pure matching effect? To do this we need to be able to break down the vector of estimates of the worker-firm fixed effect into three separate parts: one that picks up the firm effect, another the worker effect and a (pure) matching effect that reflects the quality of the match. Pre-multiplying these estimates by $\mathbf{H_Z}$ would allow us to calculate

$$\widehat{\delta}_{\boldsymbol{\varphi}} = \widehat{\delta}_{\boldsymbol{\alpha}} + \widehat{\delta}_{\boldsymbol{\theta}} + \widehat{\delta}_{\boldsymbol{\zeta}} \quad . \tag{10}$$

Unfortunately, there is no unique way to decompose the worker-firm fixed effect. One approach due to Woodcock (2015) is to impose the restriction that the pure matching effects are orthogonal to firm and worker fixed effects. This amounts to running a regression of the estimated worker-firm fixed effect on two fixed-effects —one for the firm and another for the worker. The estimates of these fixed effects give us the separate contribution of firms and workers while the residual can only be attributed to pure matching effects. With this approach, we are ascribing as much as possible of the variation on $\mathbf{M}\widehat{\boldsymbol{\varphi}}$ to the additive effects of firms and workers. This in turn means that the estimate we obtain for the firm-worker pure matching effect should be interpreted as a lower bound estimate.

7.3. Decomposing returns to education

We now quantify the importance of the different channels shaping the returns to education previously estimated. Table 7 reports the OLS results and Table 8 presents our preferred specification using an IV estimation. As before, and in both cases, column (1) presents the coefficient estimates of the benchmark results on returns to education. Column (2) shows the coefficient of the full specification. Columns (3) and (4) of Panel A report the Gelbach decomposition. Having estimated two sets of fixed effects —match and occupation —we disentangle the contribution of each fixed effects on the returns to education. In Panel B, we further decompose the match fixed effect into three components: worker, firm, and match quality.

Focusing only on the IV results, occupation heterogeneity accounts only for 5% of the returns to education. Consequently, a remarkable 95% of the return to education operates via the match heterogeneity. In other words, this decomposition shows that the match effect is responsible for 8.91 log points out of the 9.34 log points return to education: 5.94 log points accounted for by the individual component and 2.89 log points accounted for by the firm. Therefore, 31 percent of the returns operate via the allocation of workers to firms, reflecting the existence of sorting of educational levels across firms. In turn, 64 percent of the overall return to education are immune to the allocation of individuals into firms.

Relying on the procedure described in section 7.2, we obtain a lower bound for the value of the interaction worker/firm effect, the so-called match quality. This estimated effect is rather modest, contributing a mere 1 percent to the overall return to education.[16]

16. Table 7 reports a comparable decomposition, using the OLS regression and we find very similar results. Under the OLS, the contribution of the match quality component slightly increases, while the firm component decreases and the worker component remains the same.

Panel A - Gelbach Decomposition of the Return to Education

Benchmark Regression	Full Specification	Decomposition into: Match FE	Occupation FE
(1)	(2)	(3)	(4)
0.0854	0.0037	0.0772	0.0045
(0.00006)	(0.00008)	(0.00005)	(0.00020)

Panel B - Decomposition of the Match FE

Match FE	Worker	Firm	Match Quality
(1)	(2)	(3)	(4)
0.0771	0.0551	0.0219	0.0001
(0.00005)	(0.00005)	(0.00086)	(0.00003)

Table 7. Conditional Decomposition of the Return to Education - OLS (including match and occupation)

Notes: Panel A: The conditional decomposition of the return to education is based on Gelbach (2016). Column (1) reports the coefficient of the benchmark result on return to education of the OLS approach. Column (2) reports the coefficient of the full specification after including match (worker/firm) and occupation/year fixed effects. The results of the decomposition are reported in Columns (3) and (4). Panel B: The conditional decomposition of the contribution of the match(worker/firm) FE to the return to education into: the worker specific effect, the firm specific effect, and their interaction (match quality). Column (1) reproduces the coefficient of the match (worker/firm) FE contribution reported in panel A. The results of the decomposition are reported in Columns (2) to (4). Adding up the results of Columns (2) to (4) we obtain the result in Column (1). Standard errors are clustered at the corresponding group level (match or occupation, and worker or firm).

Panel A - Gelbach Decomposition of the Return to Education

Benchmark Regression	Full Specification	Decomposition into: Match FE	Occupation FE
(1)	(2)	(3)	(4)
0.0934	0.0000	0.0891	0.0043
(0.00059)	(-)	(0.00053)	(0.00070)

Panel B - Decomposition of the Match FE

Match FE	Worker	Firm	Match Quality
(1)	(2)	(3)	(4)
0.0891	0.0594	0.0289	0.0008
(0.00053)	(0.00046)	(0.00165)	(0.00015)

Table 8. Conditional Decomposition of the Return to Education - IV (including match and occupation)

Notes: Panel A: The conditional decomposition of the return to education is based on Gelbach (2016). Column (1) reports the coefficient of the benchmark result on return to education of the IV approach. Column (2) reports the coefficient of the full specification after including match (worker/firm) and occupation/year fixed effects. The results of the decomposition are reported in Columns (3) and (4). Panel B: The conditional decomposition of the contribution of the match(worker/firm) FE to the return to education into: the worker specific effect, the firm specific effect, and their interaction (match quality). Column (1) reproduces the coefficient of the match (worker/firm) FE contribution reported in panel A. The results of the decomposition are reported in Columns (2) to (4). Adding up the results of Columns (2) to (4) we obtain the result in Column (1). Standard errors are clustered at the corresponding group level (match, or occupation, and worker, or firm).

8. Conclusion

We explore the sources of the returns to education, unveiling the impact of the individual, firm, occupation, and match channels. We thereby contribute to the intersection of two strands of the literature: the role of the firm shaping the wage distribution, and the returns to education. We combine longitudinal linked employer-employee data of remarkable quality with state of the art empirical methods to address common problems in the estimation of the returns to education, namely: selection issues, and common measurement errors and confounding factors.

Schooling grants access to better paying firms and occupations. The first part of our analysis concentrates on the returns to education and to the role of firm, exploiting exogenous variation in schooling driven by changes in the institutional setting. It reveals that almost one third of the overall return on a year of education operates through the firm channel, whereas the remaining is attributable to the worker component. In the second part of the analysis we show that occupation is responsible for 12 percent, firm for 30 percent, and worker for 58 percent. In the final part of the analysis, we provide evidence that the effect of the so-called match quality in shaping the returns to education is rather modest. Overall, our results stress the importance of access to firms, and occupations, shaping the wage distribution along a dimension —returns to education —not previously explored in a comprehensive way in the literature.

In general, the practitioner should avoid an empirical strategy based on a sequential addition of covariates. "Sequential addition can obscure, overstate, or understate the true part of δ that can be given to any set of X_2 variables" (Gelbach 2016, page 530). The reader will recognize that, in this paper, we did not strictly comply with this guideline. The main reason why we present three alternative full models is because we believe that our first two AKM-type specifications (worker-firm, worker-firm-occupation) can serve as useful benchmarks for practitioners that employ more conventional specifications or have limited access to data.

References

Abowd, John M, Robert H Creecy, and Francis Kramarz (2002). "Computing person and firm effects using linked longitudinal employer-employee data." Tech. rep., Center for Economic Studies, US Census Bureau.

Abowd, John M, Francis Kramarz, and David N Margolis (1999). "High wage workers and high wage firms." *Econometrica*, 67(2), 251–333.

Acemoglu, Daron and Joshua Angrist (2000). "How large are human-capital externalities? Evidence from compulsory schooling laws." *NBER Macroeconomics Annual*, 15, 9–59.

Addario, Sabrina Di, Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten (2022). "It ain't where you're from, it's where you're at: Hiring origins, firm heterogeneity, and wages." *Journal of Econometrics*.

Addison, John, Pedro Portugal, and Hugo de Almeida Vilares (2022). "Union membership density and wages: The role of worker, firm, and job-title heterogeneity." *Journal of Econometrics*.

Andrews, M., L. Gill, T. Schank, and R. Upward (2008). "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3), 673–697.

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek (1999). "A review of estimates of the schooling/earnings relationship, with tests for publication bias." *Labour Economics*, 6(4), 453–470.

Belzil, Christian (2007). "The return to schooling in structural dynamic models: A survey." *European Economic Review*, 51(5), 1059–1105.

Blau, Francine D (1977). *Equal pay in the office*. Lexington Books.

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi (2005). "Evaluating the effect of education on earnings: Models, methods and results from the National Child Development Survey." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(3), 473–512.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler (2023). "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*.

Card, David (1999). "The causal effect of education on earnings." In *Handbook of Labor Economics*, vol. 3, edited by O. Ashenfelter and D. Card, pp. 1801–1863. Elsevier.

Card, David (2001). "Estimating the return to schooling: Progress on some persistent econometric problems." *Econometrica*, 69(5), 1127–1160.

Card, David and Ana Rute Cardoso (2022). "Wage flexibility under sectoral bargaining." *Journal of the European Economic Association*, 20(5), 2013–2061.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline (2018). "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics*, 36(S1), S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women." *The Quarterly Journal of Economics*, 131(2), 633–686.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly Journal of Economics*, 128(3), 967–1015.

Cardoso, Ana Rute, Paulo Guimarães, and Pedro Portugal (2016). "What drives the gender wage gap? A look at the role of firm and job-title heterogeneity." *Oxford Economic Papers*, 68(2), 506–524.

Cardoso, Ana Rute and Pedro Portugal (2005). "Contractual wages and the wage cushion under different bargaining settings." *Journal of Labor economics*, 23(4), 875–902.

Carneiro, Anabela, Paulo Guimarães, and Pedro Portugal (2012). "Real wages and the business cycle: Accounting for worker, firm, and job title heterogeneity." *American Economic Journal: Macroeconomics*, 4(2), 133–52.

Carneiro, Anabela, Pedro Portugal, Pedro Raposo, and Paulo M.M. Rodrigues (2023). "The persistence of wages." *Journal of Econometrics*, 233(2), 596–611.

Deming, David (2022). "Four Facts about Human Capital." *Journal of Economic Perspectives*, 36(3), 75–102.

Dhaene, Geert and Koen Jochmans (2015). "Split-panel Jackknife Estimation of Fixed-effect Models." *The Review of Economic Studies*, 82(3), 991–1030.

Eeckhout, Jan (2018). "Sorting in the labor market." *Annual Review of Economics*, 10(1), 1–29.

Eeckhout, Jan and Philipp Kircher (2011). "Identifying sorting—in theory." *The Review of Economic Studies*, 78(3), 872–906.

Engbom, Niklas and Christian Moser (2017). "Returns to education through access to higher-paying firms: Evidence from US matched employer-employee data." *American Economic Review*, 107(5), 374–78.

Gelbach, Jonah B (2016). "When do covariates matter? And which ones, and how much?" *Journal of Labor Economics*, 34(2), 509–543.

Goldschmidt, Deborah and Johannes F. Schmieder (2017). "The rise of domestic outsourcing and the evolution of the German wage structure." *The Quarterly Journal of Economics*, 132(3), 1165–1217.

Groshen, Erica L (1991a). "Sources of intra-industry wage dispersion: How much do employers matter?" *The Quarterly Journal of Economics*, 106(3), 869–884.

Groshen, Erica L (1991b). "The structure of the female/male wage differential: Is it who you are, what you do, or where you work?" *Journal of Human Resources*, pp. 457–472.

Gruetter, Max and Rafael Lalive (2009). "The importance of firms in wage determination." *Labour Economics*, 16(2), 149–160.

Guimarães, Paulo and Pedro Portugal (2010). "A simple feasible procedure to fit models with high-dimensional fixed effects." *The Stata Journal*, 10(4), 628–649.

Harmon, Colm, Hessel Oosterbeek, and Ian Walker (2003). "The returns to education: Microeconomics." *Journal of Economic Surveys*, 17(2), 115–156.

Kline, P., R. Saggio, and M. Solvsten (2020). "Leave-out Estimation of Variance Components." *Econometrica*, 88(5), 1859–1898.

Lester, Richard A (1952). "A range theory of wage differentials." *ILR Review*, 5(4), 483–500.

Lopes de Melo, Rafael (2018). "Firm wage differentials and labor market sorting: Reconciling theory and evidence." *Journal of Political Economy*, 126(1), 313–346.

Machin, Stephen and Alan Manning (2004). "A test of competitive labor market theory: The wage structure among care assistants in the south of England." *ILR Review*, 57(3), 371–385.

Macis, Mario and Fabiano Schivardi (2016). "Exports and wages: Rent sharing, workforce composition, or returns to skills?" *Journal of Labor Economics*, 34(4), 945–978.

Meng, Xin and Dominique Meurs (2004). "The gender earnings gap: Effects of institutions and firms —a comparative study of French and Australian private firms." *Oxford Economic Papers*, 56(2), 189–208.

Montenegro, Claudio E and Harry A Patrinos (2014). "Comparable estimates of returns to schooling around the world." Working Paper 7020, World Bank.

Reynolds, Lloyd George (1951). *The structure of labor markets: Wages and labor mobility in theory and practice*. Westport, Conn: Greenwood Press.

Topel, Robert H. and Michael P. Ward (1992). "Job mobility and the careers of young men." *The Quarterly Journal of Economics*, 107(2), 439–79.

Torres, Sónia, Pedro Portugal, John T Addison, and Paulo Guimarães (2018). "The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model." *Labour Economics*, 54, 47–60.

Trostel, Philip, Ian Walker, and Paul Woolley (2002). "Estimates of the economic return to schooling for 28 countries." *Labour economics*, 9(1), 1–16.

Woodcock, Simon D (2015). "Match effects." *Research in Economics*, 69(1), 100–121.

Appendix to "What lies behind returns to schooling: the role of labor market sorting and worker heterogeneity"

Appendix A.1 - Tables and Figures

	(1)
Log wages	1.7310
	(0.5731)
Age	38.7550
	(10.8705)
Tenure	8.3067
	(8.7251)
Gender (Female=1)	0.4401
	-
Schooling	9.1450
	(4.0588)
N	47,569,720

This table reports the summary statistics from Quadros de Pessoal (1995-2021).

Table A.1.1. Summary Statistics

	(1)
Age	0.1245
	(0.0007)
Age squared	-0.0033
	(0.0000)
Tenure	0.0001
	(0.0004)
Tenure squared	0.0001
	(0.0000)
Gender (Female=1)	0.7719
	(0.0041)
Instrument:	
Males born between 1929 and 1948	-
	.
Females born between 1929 and 1952	-1.4437
	(0.0112)
Females born between 1953 and 1956	-1.1909
	(0.0105)
Males born between 1949 and 1956)	-2.2167
	(0.0142)
Males and Females born between 1957 and 1979	-1.6861
	(0.0112)
Males and Females born between 1980 and 1998	-1.9348
	(0.0158)
Males and Females born after 1998	-3.6206
	(0.0196)
Time effects (μ_t)	✓
F-Test (instrument)	13,702.74
	(0.0000)
N	47,569,720
R Squared	0.2264

Table A.1.2. First stage regression for schooling

Notes: The table reports the first stage using as instrument the dummy variables, which makes the correspondence between the compulsory schooling rules with the birth date of each individual. The omitted category corresponds to males and females born between 1980 and 1998. The value of the conventional F-test for the statistical significance of the instruments is 13,702.74. The Wald F statistic Cragg-Donald weak identification test is 53,000. The Stock-Yogo weak ID critical values for 5% maximal IV relative bias is 19.28.

Panel A - Variance Decomposition			
worker	0.1830	0.1657	0.1781
firm	0.0934	0.0914	0.0932
Occupation	-	0.0224	0.0107
Match	-	-	0.2823
Match Quality	-	-	0.0119
Xβ	0.0145	0.0128	0.0154
Residual	0.0376	0.0361	0.0246
Panel B - Correlations			
$\rho(w,f)$	0.2314	0.2248	0.2596
$\rho(w,o)$	-	0.3280	0.2287
$\rho(f,o)$	-	0.1767	0.0994
$\rho(m,o)$	-	-	0.2251
Panel C - Fixed Effect Heterogeneity			
σ_w	0.5228	0.4913	0.4918
σ_f	0.2507	0.2471	0.2530
σ_o	-	0.0775	0.0546
σ_m	-	-	0.6183
σ_{mq}	-	-	0.1089

Table A.1.3. Statistical Moments from Wage Distribution

Note: The statistics are computed from the estimates from Tables 1, 4, and 6. Panel A gives the variance decomposition according to the covariances between wages and the components of the wage equation. Panel B shows the correlations between the worker, firm, occupation, and match fixed effects. Panel C provides the standard deviations of worker, firm, occupation, and match fixed effects.

	Uncorrected	Andrews et al (2008)	Kline et al (2020)
Fixed Effect Heterogeneity			
σ_w	0.5228	0.5153	0.5239
σ_f	0.2507	0.2456	0.2413
Correlation			
$\rho(w,f)$	0.2314	0.2540	0.2634

Table A.1.4. Bias Corrected estimates for the second moments of the wage distribution components

Note: The statistics are computed from the specification with worker and firm fixed effects. We provide the standard deviations of worker, and firm fixed effects, and the correlation between worker and firm fixed effects.

Working Papers

2021

1|21 Optimal Social Insurance: Insights from a Continuous-Time Stochastic Setup
João Amador | Pedro G. Rodrigues

2|21 Multivariate Fractional Integration Tests allowing for Conditional Heteroskedasticity withan Application to Return Volatility and Trading
Marina Balboa | Paulo M. M. Rodrigues | Antonio Rubia | A. M. Robert Taylor

3|21 The Role of Macroprudential Policy in Times of Trouble
Jagjit S. Chadha | Germana Corrado | Luisa Corrado | Ivan De Lorenzo Buratta

4|21 Extensions to IVX Methodsnof Inference for Return Predictability
Matei Demetrescu | Iliyan Georgiev | Paulo M. M. Rodrigues | A.M. Robert Taylor

5|21 Spectral decomposition of the information about latent variables in dynamic macroeconomic models
Nikolay Iskrev

6|21 Institutional Arrangements and Inflation Bias: A Dynamic Heterogeneous Panel Approach
Vasco Gabriel | Ioannis Lazopoulos | Diana Lima

7|21 Assessment of the effectiveness of the macroprudential measures implemented in the context of the Covid-19 pandemic
Lucas Avezum | Vítor Oliveiral | Diogo Serra

8|21 Risk shocks, due loans, and policy options: When less is more!
Paulo Júlio | José R. Maria | Sílvia Santos

9|21 Sovereign-Bank Diabolic Loop: The Government Procurement Channel!
Diana Bonfim | Miguel A. Ferreira | Francisco Queiró | Sujiao Zhao

10|21 Assessing the effectiveness of the Portuguese borrower-based measure in the Covid-19 context
Katja Neugebauer | Vítor Oliveira | Ângelo Ramos

11|21 Scrapping, Renewable Technology Adoption, and Growth
Bernardino Adão | Borghan Narajabad | Ted Temzelides

12|21 The Persistence of Wages
Anabela Carneiro | Pedro Portugal | Pedro Raposo | Paulo M.M. Rodrigues

13|21 Serial Entrepreneurs, the Macroeconomy and top income inequality
Sónia Félix | Sudipto Karmakar | Petr Sedláček

14|21 COVID-19, Lockdowns and International Trade: Evidence from Firm-Level Data
João Amador | Carlos Melo Gouveia | Ana Catarina Pimenta

15|21 The sensitivity of SME's investment and employment to the cost of debt financing
Diana Bonfim | Cláudia Custódio | Clara Raposo

16|21 The impact of a macroprudential borrower based measure on households' leverage and housing choices
Daniel Abreu | Sónia Félix | Vítor Oliveira | Fátima Silva

17|21 Permanent and temporary monetary policy shocks and the dynamics of exchange rates
Alexandre Carvalho | João Valle e Azevedo | Pedro Pires Ribeiro

18|21 On the Cleansing Effect of Recessions and Government Policy: Evidence from Covid-19
Nicholas Kozeniauskas | Pedro Moreira | Cezar Santos

19|21 Trade, Misallocation, and Capital Market Integration
Laszlo Tetenyi

20|21 Not All Shocks Are Created Equal: Assessing Heterogeneity in the Bank Lending Channel
Laura Blattner | Luísa Farinha | Gil Nogueira

21|21 Coworker Networks and the Labor Market Outcomes of Displaced Workers: Evidence from Portugal
Jose Garcia-Louzao | Marta Silva

22|21 Markups and Financial Shocks
Philipp Meinen | Ana Cristina Soares

2022

1|22 Business cycle clocks: Time to get circular
Nuno Lourenço | António Rua

2|22 The Augmented Bank Balance-Sheet Channel of Monetary Policy
Christian Bittner | Diana Bonfim | Florian Heider | Farzad Saidi | Glenn Schepens | Carla Soares

3|22 Optimal cooperative taxation in the global economy
V. V. Chari | Juan Pablo Nicolini | Pedro Teles

4|22 How Bad Can Financial Crises Be? A GDP Tail Risk Assessment for Portugal
Ivan De Lorenzo Buratta | Marina Feliciano | Duarte Maia

5|22 Comparing estimated structural models of different complexities: What do we learn?
Paulo Júlio | José R. Maria

6|22 Survival of the fittest: Tourism Exposure and Firm Survival
Filipe B. Caires | Hugo Reis | Paulo M. M. Rodrigues

7|22 Mind the Build-up: Quantifying Tail Risks for Credit Growth in Portugal
Ivan de Lorenzo Buratta | Marina Feliciano | Duarte Maia

8|22 Forgetting Approaches to Improve Forecasting
Robert Hill | Paulo M. M. Rodrigues

9|22 Determinants of Cost of Equity for listed euro area banks
Gabriel Zsurkis

10|22 Real effects of imperfect bank-firm matching
Luísa Farinha | Sotirios Kokas | Enrico Sette | Serafeim Tsoukas

11|22 The solvency and funding cost nexus – the role of market stigma for buffer usability
Helena Carvalho | Lucas Avezum | Fátima Silva

12|22 Stayin' alive? Government support measures in Portugal during the Covid-19 pandemic
Márcio Mateus | Katja Neugebauer

13|22 Cross-Sectional Error Dependence in Panel Quantile Regressions
Matei Demetrescu | Mehdi Hosseinkouchack | Paulo M. M. Rodrigues

14|22 Multinationals and services imports from havens: when policies stand in the way of tax planning
Joana Garcia

15|22 Identification and Estimation of Continuous-Time Job Search Models with Preference Shocks
Peter Arcidiacono | Attila Gyetvai | Arnaud Maurel | Ekaterina Jardim

2023

14|23 Prudential policy treatments to the COVID-19 economic crisis: an assessment of the effects.
Ivan De Lorenzo Buratta | Diana Lima | Duarte Maia

15|23 Structural and cyclical capital instruments in the 3D model: a simulation for Portugal
Diana Lima | Duarte Maia | Ana Pereira

16|23 Currency choices and the role of the U.S. dollar in international services trade
João Amador | Joana Garcia

17|23 Navigating with a compass: Charting the course of underlying inflation
Nuno Lourenço | João Quelhas | António Rua

18|23 The determinants of the loss given default of residential mortgage loans in Portugal
Márcio Mateus | Ângelo Ramos

19|23 Regulating Artificial Intelligence
João Guerreiro | Sérgio Rebelo | Pedro Teles

20|23 The Impact of a Higher Cost of Credit on Exporters: Evidence from a Change in Banking Regulation
João Monteiro | Pedro Dias Moreira

21|23 Taxes and Migration Flows: Preferential Tax Schemes for High-Skill Immigrants
João Brogueira de Sousa | Pedro Teles

22|23 What lies behind returns to schooling: the role of labor market sorting and worker heterogeneity
Pedro Portugal | Hugo Reis | Paulo Guimarães
Ana Rute Cardoso;