Wage inequality, business strategy and productivity: evidence from Portugal[*]

P. Ferreira[†], J. Cerejeira, M. Portela, S. Sousa

Department of Economics, University of Minho[‡]

Preliminary version: please do not quote

3rd November 2014

Abstract

Using longitudinal linked employer-employee data we estimate firm- and worker-effects in wages using AKM methods of analysis. We then select different groups of firms, associated to different productivity levels and HRM profiles, and identify differences across the groups both in terms the observed composition of the workforce and in terms of the estimated unobserved factors. Additionaly, we decompose wage inequality by the subgroups of firms. Between group inequality explains a very small fraction of wage inequality. We follow up and analyse the contribution of the estimated effects of observed and unobserved characteristics of workers and firms to within group wage inequality. We conclude that time invariant unmeasured human capital is the major source of inequality within groups of firms, yet differences in compensation policies across firms (net of industry effects) are also a relevant factor.

Keywords: Wage structure, human capital, productivity, HRM, LEED

JEL Classification: C33, J31, J53

[*]**Acknowledgments:** This research was partially funded by Fundação para a Ciência e a Tecnologia (Portugal) through the research project ref. no. PTDC/EGE-ECO/122126/2010 and the Applied Microeconomics Research Unit (NIMA), award no. PEst-OE/EGE/UI3181/2011.

[†]Corresponding author: Dept. of Economics, EEG – University of Minho, Braga 4710-057, Portugal. Email: priscila@eeg.uminho.pt, Phone: +351253601938, Fax: +351253601380

[‡]Ferreira is affiliated to the research unit NIMA; Cerejeira, Portela and Sousa are affiliated to the research unit NIPE.

1 Introduction

Significant inter-industry wage differentials are commonly found in most empirical analyses of wages using cross sectional data. The wage structure across industries can arise as a consequence of imperfect measures of worker and firm characteristics (Krueger and Summers, 1988). That is, industries may differ on the type of workers employed or on the compensation policies of firms. These differences may be responsible for the wage structure, e.g., high wage industries can either be hiring high wage workers or be composed of high wage firms, or have a combination of both. Previous research has attempted to decompose wage differentials paying particular attention to the structure (observed characteristics) of firms (Davis and Haltiwanger, 1991); others relate wages to different supervision schemes across industries and finds no evidence that interindustry differences in monitoring contribute to inter-industry wage differentials which contradicts explanations of industry wage premiums being caused by efficiency wage models of shirking (Neal, 1993). Groshen (1991) studies intraindustry establishment wage differentials and concludes that they are not random variations or returns to usual measures of human capital and suggests that further investigation is needed into inter-industry wage differences on sorting by unmeasured worker quality, compensating variations, efficiency wages, and rent-sharing.

Since the wage structure can be a consequence of omitted variable bias in cross-sectional analyses, some authors attempt to identify the sources of the industry wage structure using longitudinal linked employer-employee data. For example, Goux and Maurin (1999) use French data and find that the wage structure is mainly due to unmeasured labour quality and that the potential wage gains from switching industries would be less than 3%. Also using French data, Abowd et al. [AKM] (1999) conclude that person effects are relatively more important in explaining the differentials found in cross-sectional analyses. The same result is obtained by Abowd, Finer and Kramarz (1999) with data for the State of Washington and applying the same decomposition as AKM (1999). Controlling for worker-firm match effects Woodcock (2008)[1] using American data finds that firm effects are responsible for 72% of the variance in raw interindustry wage differentials, and Ferreira (2009) using Portuguese data finds that 27% of the cross-sectional wage differences across industries are due to person unmeasured effects and 70% of the wage structure are generated by firm effects.[2] Overall, these studies provide

[1] In contrast with the other studies using linked employer-employee that disaggregate industry coding to a detailed level (more than 90 industries) Woodcock (2008) uses only 8 SIC Major Divisions.

[2] The difference between the findings of Goux and Maurin (1999) and the other papers using matched

evidence that, after controlling for worker, firm and match unobserved characteristics, (true) inter-industry wage differentials exist and persist over time. An industry can be regarded as a wage contour and the industry wage effect is determined by the pay policies of firms that compose the contour. However, empirical evidence suggests that firms are heterogeneous and that they differ with a high degree of persistence in several respects, e.g., human resource management, productivity, and pay policies; while workers differ in education, physical attributes and ability and human capital even within narrowly defined industries (Doms et al., 1997; Syverson, 2011).

In this paper we use the Portuguese LEED *Quadros de Pessoal* to, firstly, estimate unobserved person- and firm-effects in wages using the AKM (1999) method. This allows us to identify high- and low-wage firms, and the industry wage structure. Secondly, we analyse differences and similarities of high-road and low-road industries and firms with respect to elements of human resources strategy. We investigate whether different management structures, usually associated to higher-levels of productivity, are associated to different pay policies of firms (high- or low-wage firms) within an industry. Following Tor (2012), we identify different management strategies of firms by focusing on contrasts between: (i) exporters versus non-exporters; (ii) multi- vs single-plant firms; (iii) foreign- versus domestic-owned firms.

Our results indicate the existence of a different wage structure across industries, after controlling for time-invariant unobserved characteristics of the workers, which suggests that the differentials estimated from cross-sectional data are not due to compositional effects of the workforce and that different firms pay different wages to workers with the same characteristics (measured or unmeasured). This means that substantial wage growth can be attained via mobility across industries. We also find considerable dispersion in firm compensation policies within both high- and low-road industries. Given this heterogeneity in firm compensation policies within industries, and assuming that differences in productivity lead to differences in wages, we followed up to describe and decompose wage inequality into some of its components over mutually exclusive groups of firms defined according to productivity levels. Our results suggest there are differences across the groups of firms both in terms the observed composition of the workforce and in terms of the estimated unobserved factors. The results also indicate that between group inequality explains a very small fraction of wage inequality, i.e., it is not the difference in

worker-firm data might be essentially due to the different techniques applied. The other studies use similar statistical methods to decompose the raw interindustry wage differentials found in cross-sectional data.

productivity levels that generates the wage structure. The wage structure is mostly due to some characteristics that are present in every group of firms. When we analyse the contribution of the estimated effects of observed and unobserved characteristics of workers and firms to within group wage inequality we conclude that time invariant unmeasured human capital is the major source of inequality within groups of firms, yet differences in compensation policies across firms (net of industry effects) are also a relevant factor. Therefore, it is by decreasing the differences in the workforce human capital (transversal skills, training, specific capital), that wage differences would be reduced.

The paper is structured as follows. In the next section we describe the data used and the statistical methods of the analysis. In section 3 we present and discuss our results. Firstly we identify which firms are high-wage and which are low wage and the industry wage structure. Secondly, we describe differences, both in terms of the observed composition of the workforce and in terms of the estimated unobserved factors, across subgroups of firm profiles. Thirdly, we describe wage inequality and decompose it (in terms of between- and within-group inequality) by firm type. We follow up and analyse the contribution of the estimated unobserved firm and worker effects to within group wage inequality. Section 4 summarizes and concludes our analysis.

2 Data and methods

2.1 Data

The main data set used in this analysis are the Quadros de Pessoal (QP) from Portugal, a longitudinal data set with matched information on workers and firms. These data have been collected annually since 1985 by the Portuguese Ministry of Employment and the participation of firms with registered employees is compulsory. The data include all firms (over 250 thousand per year) and employees (more than two million per year) within the Portuguese private sector. Each firm and each worker have a unique registration number which allows them to be traced over time. All information (on workers, plants, and firms) is reported by the firm and, generally, relates to the situation observed in the month when the survey is collected. The administrative nature of the data and its public availability at the workplace, as determined by law, result in a high degree of coverage and reliability. Information on workers includes, for example, gender, age, educational level, qualifications (skill), occupation, type of contract of employment, date of entry to the firm and date of last promotion, monthly hours of work and earnings (split into a few components). Firm level data include, for example, the year of creation, industry, location, total number of workers, number of establishments, sales volume, legal structure and ownership structure (shares of participation of private, public or foreign capital). We also use the International Trade data collected by the Portuguese Office of National Statiscs (INE). This data set includes the universe of monthly export and import transactions by firms located in Portugal. Each transaction includes the firm's identifier making it possible combine this data with the QP data. For the purposes of our analysis we use the data on export transactions to identify whether or not the firm is an exporter.

In our empirical analyses we estimate models with two-way fixed effects. Abowd et al. (2002) show that the identification of the unobserved effects using fixed effects techniques can be obtained by constructing groups of connected workers and firms. Mobility of workers across firms is necessary for constructing such groups, each of which contains all workers who have ever worked for any particular firm and all the firms at which any particular worker was ever employed. We restrict our analysis to the biggest connected group of workers and firms for the years 2002 to 2009, and focus workers employed in firms operating within industries from the manufacturing and services sectors for which there is non-missing information in the covariates relevant to our study. Our sample contains 13,385,663 worker-year observations, relating to 3,323,016 workers and 246,564 firms.[3]

Our dependent variable is the log hourly real wage (in euro) of the worker and is constructed by adding up its components which are: (i) the base pay – gross amount of money paid (in the reference month) to workers on a regular monthly basis for the normal hours of work; (ii) tenure related payments; and (iii) regular payments. To obtain the real hourly wages, the monthly wages are deflated using the CPI and then divided by the number of hours of work. The vector of observed time-varying covariates includes: age and seniority at the firm (and their squares), the type of contract of employment (whether open- or closed-end contract), whether the worker works part- or full-time, the skill level (low-, medium, high-skill), and the education level (ISCED1, ISCED2, ISCED3, and ISCED5/6)[4] We also control for firm characteristics, such as: log of the

[3] Abowd et al. (2002) show that the identification of the unobserved effects using fixed effects techniques can be obtained by constructing groups of connected workers and firms. Mobility of workers across firms is necessary for constructing such groups, each of which contains all workers who have ever worked for any particular firm and all the firms at which any particular worker was ever employed. Since 93% of our original sample is part of the biggest connected group (the sample we work with), we conclude that there is considerable worker mobility in the economy.

[4] ISCED 1 – up to primary education, includes the first and second stages of basic education in

firm size (size is measured by the number of workers employed by the firm), the log of real sales volume (in thousands of euro), ownership status (whether private, public or foreign owned, depending on whether more than 50% of the firms' social capital is owned by private, public or foreign investors, respectively), whether the firm is an exporter and whether the firm is multi-plant. Industry, region (NUTs I: North, Center, Lisbon and Tejo Valley, Inland, Algarve, and Islands) and year dummies (2002-2009) are also included to control for regional effects and aggregate economic shocks. Descriptive statistics of our sample are presented on Table 1.

[Table 1 about here]

Over the period the average hourly wage was 4.57 euro (hence its natural log is 1.52). 57% of the workers in our sample are males, the average age is 37 years and the mean seniority at the firm is nearly 8 years. Considering education and skill, 68% of the workers have only up to 9 years os schooling (ISCED1 and 2), and about 11% of the workforce has a university degree. The distribution of workers by skill levels is more even, 41% of them are medium skilled and 21% high-skilled. As expected, a larger proportion of workers are in open-ended contracts (73%) and in full-time jobs (97%). From these statistics it seems that firms in our sample prefer temporary-contracts (27%) to part-time jobs (3%). Regarding the observed characteristics of firms, the weighted mean of the real sales volume of firms in our sample is about 4 million euro (hence the natural log is 15,22) and firms employ, on average, 77 workers (ln firm size is 4,34). About 47% of the workers are on firms that export, 38% in multi-establishment firms and 12% of them in firms that are owned by foreigners. The majority of our workers are located in the Lisbon and Tejo Valley (42%) and in the North Coast (28%) regions. Worker-year observations are fairly evenly distributed across the years under analysis (proportions ranging between 11 and 14%). We use a fine disaggregation of industry classification and consider 46 distinct industries. These are listed in Table 2 together with their relative frequencies in the data.

[Table 2 about here]

As we have mentioned before, we focus on two sectors of activity: manufacturing and services. Looking at the distribution of the workers over these two sectors 71% of the

Portugal (up to 6 years of schooling); ISCED 2 – lower secondary education, includes the third stage of basic education (9 years of schooling); ISCED 3 – upper secondary education (12 years of schooling); ISCED 5/6 – higher education, includes first and second stages of tertiary education (more than 15 years of schooling corresponding to university degrees).

workers are employed in firms providing services while the remaining 29% are employed in manufacturing firms. The industries that have larger shares of the workforce are Construction (12,5%), Other business activities (10%), Retail (9%) and Wholesale (8%) trade, and Hotels and restaurants (7%).

2.2 Empirical strategy

In the competitive model the wages of workers do not depend on firm or industry affiliation. This prediction is usually tested by defining a wage function as follows:

$$y_{ijt} = x_{\cdot t}\beta + k_{(j(i,t))}\kappa + \epsilon_{ij} \tag{1}$$

where y_{ijt} is the logarithm of real hourly wages of worker $i = 1, ..., N$ in firm $j = 1, ..., J$ in period $t = 1, ..., T$; $x_{\cdot t}$ is the vector of observed time varying covariates of workers (i) and firms (j); $k = 1, ..., K$ is a vector of mutually exclusive dummy variables indicating the industry affiliation of firm j; and ϵ_{ij} is the idiosyncratic error. If wages do not depend on industry affiliation, then the parameters κ should be jointly equal to zero. We test this prediction in the next section.

If cross sectional industry effects are statistically different from zero, raw-interindustry wage differentials exist. We proceed by investigating if these effects hold after controlling for unobserved effects of workers and firms. To this end we specify a competitive model of wage determination that considers the longitudinal structure of the data and includes observed characteristics of workers and firms, as well as time invariant unobserved worker and firm characteristics as determinants of wages.[5] The estimation of this model allows for the identification of "true" interindustry wage structure, that is, wage differentials that exist after controlling for the unobserved / unmeasured quality of the workers and compensation policies of firms. This structure is given by the average of the firm effects within industries.

The person and firm effects model estimates a wage equation of the type:

$$y = X\beta + D\theta + F\psi + \xi \tag{2}$$

where y is a $(N^* \times 1)$ vector of log monthly real wages (in deviations from the grand mean); $X_{(N^* \times Z)}$ is the matrix of observable time varying covariates (also in deviations from the grand mean); $D_{(N^* \times N)}$ is the matrix of indicators for worker $i = 1, ..., N$; $F_{(N^* \times J)}$

[5] This section follows very closely the framework developed in AKM (1999).

is the matrix of indicators for the firm at which worker i works at period t. The set of parameters to estimate are β, the $Z \times 1$ vector of coefficients on the covariates; θ, the $N \times 1$ vector of worker effects; ψ, the $J \times 1$ vector of firm effects. Worker effects capture time invariant characteristics of workers that affect the worker's wages in any firm across time, it is a portable component of compensation. Firm effects capture time invariant characteristics of firms that affect the wages of all workers within a firm in the same way across time. We can perform an orthogonal decomposition of the time invariant worker effects (θ) into observable (μ) and unobservable (α) components as follows:

$$\theta_i = \alpha_i + \mu_i \eta, \tag{3}$$

where α can be interpreted as time invariant human capital or abilities of the worker that are not measured in our survey, while η reports the effects of observed time-invariant covariates of the worker (which in our study are gender and education). A similar decomposition can be made for the time invariant firm effects. Considering that each industry is a wage contour, an industry effect can be defined as a characteristic of the firm and the pure interindustry wage structure, conditional on the same information as in equation (2), is defined as κ_k for some industry classification $k = 1, ..., K$. Being a characteristic of the firm, it follows that the definition of the pure industry effect (κ_k) is the aggregation of the pure firm effects (ψ) within the industry, that is

$$\kappa_k \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\frac{\mathbf{1}(\mathrm{K}(\mathrm{J}(i,t)) = k)\psi_{\mathrm{J}(i,t)}}{N_k} \right] \tag{4}$$

where

$$N_k \equiv \sum_{j=1}^{J} \mathbf{1}(\mathrm{K}(j) = k) N_j$$

and $\mathbf{K}(j)$ is a function denoting the industry affiliation of firm j. This aggregation of J firm effects into κ_k industry effects, weighted so as to be representative of individuals, corresponds to including industry indicator variables in equation (2), $\kappa_{\mathrm{K}(\mathrm{J}(i,t))}$, and defining what is left of the pure firm effect as a deviation from industry effects, $\psi_{J(i,t)} - \kappa_{\mathrm{K}(\mathrm{J}(i,t))}$ (firm effects net of industry effects).[6]

We will use our estimates of the pure interindustry wage structure to identify high- and

[6] Authors attempt to use industry classifications as detailed as to have more than 90 industry codes. The reason for decomposing industrial aggregates into the most detailed level possible is related to the possibility that average compensation policies of firms may vary across finer levels of classification and not within aggregates, and so estimates can be subject to aggregation biases. The pure industry effects, however, are not subject to this bias because they are computed from firm-level estimates.

low-road industries. We will present alternative definitions of high/low-road industries and describe them in terms of observed and unobserved characteristics. This allows the identification of similarities and differences in behaviour and composition depending on the business strategy. We then decompose the wage differentials across high/low roads in terms of within and between-inequality. Since our results suggest that most of the wage differences are due to within inequality (that is, wage structures regardless of the business strategy type), we decompose the within-inequality in terms of shares due to observed factors and unobserved characteristics of workers (human capital and ability) and firms (compensation policies).

3 Results

3.1 Firm compensation policies and the interindustry wage structure

To assess the existence of cross-sectional relative wages across industries we regress model (1) on the annual data for the period from 2002 to 2009. The vector X includes covariates on workers and firms as described in subsection 2.1, and k assigns each firm to each of the $K = 46$ defined industries. An overview of the distribution of the year-on-year raw interindustry wage differentials is shown in Figure 1. Our estimates suggest the existence of a cross sectional interindustry wage structure. The industry parameters are jointly statistically significant (yearly F-tests reject the hypothesis that the estimated industry coefficients are jointly equal to zero) and there is considerable dispersion on the wage premia across industries. Considerable dispersion in raw-interindustry wage differentials was also found by Hartog et al. (2000) for the period 1982-1992 in Portugal, the authors suggest these results are similar to those of countries with a decentralised wage setting environment and note the potential flexibility to exploit industry (firm) specific conditions.

[Figure 1 about here]

However, as mentioned previously, wage differentials obtained from cross-sectional data may be generated by temporary disequilibrium in the labour market (thus being a transitory phenomenon) or by systematic differences in unobserved labour quality or in firm characteristics across industries (Krueger and Summers, 1998; Gibbons and Katz, 1992). Several studies have estimated wage equations where unobserved characteristics of workers and firms were controlled for and concluded that the composition of the workforce was the main factor driving the inter-industry wage structure (Goux and Maurin, 1999; AKM, 1999; Abowd, Finer and Kramarz, 1999). To investigate the existence of a "true" interindustry wage structure, we have estimated a wage equation where we include time invariant person and firm effects (as specified in equation 2). The estimated yearly interindustry wage structure, the yearly average of firm effects within industries, for the period 2002-2009 is presented in Figure 2. From Figure 2 we conclude that, after controlling for unmeasured firm and worker characteristics, wages differ across industries and that the distribution of interindustry wage premia shows considerable dispersion. Moreover, the distribution of the genuine industry effects (Figure 2) is very similar to that of the raw interindustry wage differentials estimated from cross-sectional data (Figure 1). Both distributions show more variability in the upper tail of the wage structure than in the lower tail, this may be caused by the existence of a minimum wage law.

[Figure 2 about here]

The existence of a wage structure across industries, after controlling for time-invariant unobserved characteristics of the workers, suggests that the differentials estimated from cross-sectional data are not due to compositional effects of the workforce and that different firms pay different wages to workers with the same characteristics (measured or unmeasured). This means that substantial wage growth can be attained via mobility across industries. Rent-sharing (which relates profitability and wages) or efficiency wages (which relates specific-skills and wages) are possible explanations for the existence of such a wage structure. Both theories suggest that firms may find it profitable to pay wages above the competitive level. If firms across industries have different compensation policies, an interindustry wage structure arises. On the other hand, if firms are paying the value of the marginal product of their workforce, then high-wage industries are more productive than "low-road" industries and differences in productivity are the mechanism generating the wage differentials across industries. The dispersion in the wage premia paid by firms within industries, however, suggests that there are differences in productivity across firms within an industry.

We document the differences across firm compensation policies (firm-effects) within industries and check whether there is an association between the industry wage premia and the dispersion of premia across firms within an industry. In Figure 3 we plot the interindustry wage premia (average of firm effects that compose the wage contour across the period 2002-2009) and the standard deviation of the firm effects within the industry.

We find considerable dispersion in firm compensation policies within both high- and low-road industries suggesting that there is weak association between the industry wage premia and the dispersion of compensation policies across firms within an industry. In fact, the correlation between the industry wage premia and the dispersion of firm effects within industries is small and positive (0.14, as is reflected by the small positive slope of the linear prediction line depicted in Figure 3).

[Figure 3 about here]

Within the group of high-wage and low-dispersion industries (the median being the threshold separating high/low-wage (high/low-dispersion)) we find: 41 Water collection, treatment and distribution; 64 Post and telecommunications, 65 Financial intermediation and 66 Insurance. These are industries with low levels of competition. Amongst the industries that pay, on average, high wages and have high dispersion in the premia paid, the industries forming the wage contour include: 67Auxilliary activities to financial intermediation; 61 Water transport; 63 Travel agents and tourism; 24 Chemicals and chemical products. Within the set of low industry wage premia and low dispersion, we find low value-added manufactures such as: 17 Textiles, 18 Manufacture of wearing apparel, dressing and dying of fur; 19 Leather; and 36 Manufacture of furniture. Industries with low wages and high dispersion in firm compensation policies include: 45 Construction; 50 Sales and maintenance of vehicles; 52 Retail trade; 74 Other businesses; and 93 Other services. This group involves a series of firms that undertake various low value-added businesses and services, and the mix of activities involved may be the cause of larger dispersion in the compensation policies across firms.

3.2 Measuring and decomposing wage differentials: productivity of firms and inequality

Firms are heterogeneous and empirical evidence suggests that they persistently differ in several respects relating to, for example, their human resource management, productivity, and pay policies (Syverson, 2011). The analysis in the previous section supports this argument in that it concludes for the existence of different compensation policies across firms within industries. Assuming that differences in productivity lead to differences in wages, in this section we describe and decompose wage inequality into some of its components over mutually exclusive groups of firms.[7] Therefore, to inspect different

[7] Cowell and Fiorio (2010) provide a good survey on the decomposition approaches that we use.

behaviours of high- and low-road industries and firms we first compare the top and bottom 5 industries, the top (bottom) include the 5 industries for which the estimated effect (κ) is highest, while the bottom includes the 5 industries for which the estimated effect is lowest. We also compare the behaviour of the firms that are ranked in the top (bottom) decile of the distribution of the estimated firm effects (ψ). However, the structure and organisational form of the firm is likely to affect productivity (Garicano and Heaton 2007), as does the quality of the management practices and characteristics (Bloom and Van Reenen 2007; Bertrand and Schoar 2003). The latter however are often unobserved in surveys of firms (Syverson 2011). To account for differences in productivity we follow Tor (2012) and focus on contrasts between: (i) exporting versus non-exporting firms; ii) multi- versus single-plant firms; iii) foreign versus domestic owned firms. Exporting firms have been shown to be more productive than non-exporting firms even prior to exporting, and the onset of exporting further raises productivity (Girma et al., 2004). Evidence suggests that multi-establishment firms are likely to have accumulated more knowledge about the economic environment and may also have a more experienced management structure than single establishment firms (Audretsch and Mahmood, 1995), thus reducing their chances of failure when faced with negative economic shocks. Previous research also suggests that foreign-owned firms have higher levels of productivity than home-owned, and firms which are acquired by foreign companies exhibit an increase in labour productivity of 13% (e.g. Conyon, et al., 2002).

To measure wage inequality we adopt some alternative indicators: (i) the decile dispersion ratio (P90/P10); (ii) the Gini coefficient; and (iii) generalized entropy measures: Theil's T and L. The decile dispersion ratio is computed as the ratio of the average real hourly wage of the richest 10 percent group of workers to the average real hourly wage of the poorest 10 percent of workers. Hence the decile dispersion ratio expresses the wages of the top 10 percent (the "high-wage") as a multiple of that of workers in the poorest decile of the wage distribution (the "low-wage"). However, this ratio ignores information about incomes in the middle of the wage distribution and does not use information about the distribution of wages within the top and bottom deciles. The Gini coefficient plots the proportion of total wages that is cumulatively earned by the bottom x% of workers. The lower the Gini coefficient the more equal the distribution, with 0 corresponding to complete equality, while higher Gini coefficients indicate a more unequal distribution. For example, a Gini coefficient of 1 corresponds to one worker receiving 100% of total wages. The Gini coefficient may not be perfectly decomposable or additive across groups (when ranking by subgroup wages overlaps a positive residual is generated, this has implications for the decompositions of subsection 3.2.2). The Generalized Entropy (GE) measures have the property of additive decomposability, that is, they are decomposable without a residual term. The values of GE(α) measures vary between zero and infinity, with zero representing an equal distribution and higher values representing higher levels of inequality. The parameter α in the GE class represents the weight given to distances between wages at different parts of the wage distribution and can take any real value. For lower values of α, GE is more sensitive to changes in the lower tail of the distribution, and for higher values of α GE is more sensitive to changes that affect the upper tail. The most common values of α used are 0, 1, and 2. GE(1) corresponds to Theil's T index, GE(0) is known as Theil's L and is sometimes referred to as the mean log deviation measure.

In the sections that follow we first describe the various subgroups of the population in terms of means of the observed and estimated factors, as well as in terms of correlations between the unobserved effects of workers and firms (proxy for positive/negative sorting in the labour market). We then compute various measures of inequality for the Portuguese labour market overall and for the subgroups of the population we presented earlier in this section. We follow up and decompose the overall structure of inequality in terms of between and within group inequality and in terms of shares of contribution to inequality by factors such as the observed covariates, the estimated time invariant (measured and unmeasured) effects of workers and firms, and the unexplained part of the model.

3.2.1 Differences in composition across partitions of the population

In this subsection we describe the different groups of the population in terms: (i) of the time-varying observed characteristics of workers and firms; (ii) of the effects estimated from the model specified in equation (2); and (iii) of the pattern of sorting of workers and firms. The comparison of characteristics of workers and firms helps establishing differences and similarities across types of business strategies and assess whether wages and employment are a function of firm productivity. The relevant statistics are presented in Table 3.

[Table 3 about here]

Considering the set of observed covariates (first panel of Table 3), our results suggest that more productive industries and firms employ larger shares of highly-educated and high-skilled workers, and have lower shares of workers in temporary contracts. Moreover, they report higher sales volumes, and are on average larger firms. The average wages paid by the groups of more productive firms pay are higher than those paid by the less productive ones. Looking at the composition of the top (bottom) 5 industries (column ii) the top 5 industries pay, on average, the double of the wage paid by the low-road industries. The share of workers with university degrees in high-roads in 3 times larger than that of the economy overall (34% vs 11%) and they use twice as much high-skilled workers than that we observed for the economy-wide average (49% vs 21%). Within the top 5 industries we find larger proportions of more productive firms: exporters (54%), is multiplant (86%), and foreign owned (21%) than what we observe in the economy overall. On the other hand, low-roads employ only a very small share of highly-educated and high-skilled workers (2% and 8%, respectively), and the proportion of more productive firms (multiplant, foreign owned) is also smaller. If instead we rank firms in terms of their compensation policies (column iii) we conclude that firms in the top decile of the firm-effects distribution have lower shares of temporary workers, larger shares of educated and skilled workers, they are larger in size and have bigger sales volumes than firms in the first decile. The share of exporting, multiplant and foreign-owned firms is also larger in the top decile. When we describe our sample in terms of business strategy across industries, i.e. whether firms are exporting (column iv), multiplant (column v) or foreign owned (column vi) the same sort of differences arises. Therefore, both the high road industries and the high-road firms are composed, in terms of observed covariates, differently from their counterparts.

Inspection of the mean of the estimated effects (second panel of Table 3) reveals that the average of the effects within high-road industries and firms are above the economy-wide average. High-roads perform better in terms of the observed time-varying characteristics of workers and firms (observed covariates), but they are also composed of better workers and firms with respect to the estimated time-invariant unmeasured effects (the unobserved human capital, α, and the firm effects net of industry effects). For example, the mean of the estimated unobserved human capital for the economy as a whole is 0.06, whilst the mean in the top 5 industries is 0.35, in the top decile firms is 0.08, for exporters it is 0.09, for multiplant firms is 0.10 and for foreign owned firms is 0.11. Interestingly, the estimated firm effects (net of industry effects) are more or less similar across business types.[8] Overall, the lack of clear differences in the component of the estimated firm effects suggests that most differences across business types may be generated by differences

[8] These effects are clearly different when we split the economy into the deciles of firms, in column iii, this is explained by the splitting itself since it is done using the distribution of firm effects.

in characteristics of the workforce rather than in differences in compensation policies of firms (this will be tested in subsection 3.2.3).

Considering the differences in composition in the observed covariates and in the effects estimated from model specified in equation (2) our results are in line with those of Doms et al. (1997) in that workers in more productive firms and industries have larger observed and unobserved abilities. The existence of complementarities between workers abilities and firm productivity motivated our brief analysis on the type of matching, defined as the correlation between the time invariant effects of workers and firms, observed across the different groups of the population. There is positive assortative matching when the most productive firms employ the most skilled workers (Davidson et al., 2010). In our setting, similar to Goux and Maurin (1999) and Abowd et al. (2002), this translates into positive correlation coefficients between worker and firm estimated effects. Considering firstly the economy overall (column i) we find a very small and positive (0.09) correlation between the time-invariant worker and firm effects (θ and ψ, respectively). But these effects include the time invariant characteristics of workers (gender and education) and firms (industry affiliation). The correlation of unmeasured human capital (α) and firm compensation policies (ψ) is nearly zero, suggesting there is no assortative matching in the labour market. Looking at the type of matching across high- and low-road industries (column ii) there seems to be negative matching, that is, either we have high-wage workers in low-wage industries (bottom 5 with correlations of -0.15 and -0.20) or we have low-wage workers in high-wage industries (top 5 with correlations of -0.22 and -0.24), the same type of conclusions can be reported when we split the sample into the top and bottom decile of firm effects (column iii). Results are clearly different when our splitting of the population is based on business strategies. We find positive assortative matching in firms that export (column iv), in the multi-establishment (column v) and in the foreign owned firms (column vi), and negative matching in their counterparts. Helpman et al. (2010) argued that more productive firms screen more intensely and have a workforce of higher average ability than less productive firms, which seems to be the case for these 3 types of businesses. Our results suggest that these firms seem to have larger returns to worker-firm complementarities, this must have increased their incentive to screen workers, hence increase the matching quality.[9]

[9]Helpman (2010) and Davidson et al. (2010) also conclude that when the economy is opened to trade more productive firms select into exporting, this increases their revenue relative to less productive firms, which further increases the incentives to screen workers. Therefore exporting increases the wages paid by a firm with a given productivity. Overall, increased openness improves the efficiency of matching in the labour market (particularly whithin the group of more productive firms with more comparative

3.2.2 Inequality within and between high- and low-roads

In this subsection we compute several measures of inequality and explore the structure of inequality. We decompose the Theil's T index (GE(1)) by subgroups of the population (relating to different business strategies) to infer on the extent to which differences between the population sub-groups explain overall inequality in wages. In Table 4 we describe the distribution of the population and of total wages across the subgroups we defined previously, we also show the different measures of inequality (P90/P10 dispersion ratio, Gini, GE(0), G(1)) and decompose wage inequality (based on the Theil's T index (GE(1)) into within and between-group components. In such decomposition, the component relating to between inequality captures the share of inequality that is due to the variability of wages across the different sub-groups of the population, while the component relating to within-inequality captures the share of overall inequality that is due to the dispersion of wages within each sub-group of the population. The decomposability requirements of inequality indices (Shorrocks, 1984) lead us to focus on one of the Generalized Entropy indices, i.e., the Theil's T index (GE(1)).[10] Given the partitions of the population the most general decomposition of an inequality index has the following form

$$I = I_{within} + I_{between} + residual \tag{5}$$

The GE indices are perfectly decomposable into within and between elements without a residual term, and their decomposition can be expressed as

$$GE_\alpha = \sum_{g=1}^{m} \left(\frac{\overline{y}_g}{\overline{y}}\right)^{\alpha} \left(\frac{n_g}{n}\right)^{1-\alpha} GE(\alpha)_g + \frac{1}{\alpha^2 - \alpha}\left(\sum_{g=1}^{m} \left(\frac{\overline{y}_g}{\overline{y}}\right)^{\alpha} \frac{n_g}{n} - 1\right) \tag{6}$$

where GE_α is the overall GE index and $GE(\alpha)_g$ is the GE index of the g^{th} subgroup. The first term on the RHS is the weighted average of the GE indices for each subgroup $g = 1, ..., m$ (where weights are represented by the total income share, i.e., the product of population shares and relative mean incomes). This is the within part of the decomposition. The second term is the index calculated using the subgroup means instead of actual wages, in order to capture variability amongst the groups, and it gives the between part of the decomposition.

Our inequality indices and the decomposition of Theil's T index are presented in Table

advantages).

[10] Results on the decomposition on within and between components of inequality is similar using the other GE index, G(0). Results can be shown upon request from the author.

4. The first column of the Table shows estimates of our inequality indices for the labour market overall. The following three columns in the table report estimate of the Gini coefficient and the GE indices by sub-groups of the population according to exporting status, whether or not the firm is multi-plant and by ownership (whether national or foreign capital). The labour market's P90/P10 dispersion ratio reveals that workers in the top decile of the wage distribution receive wages that are almost four times those of workers in the bottom decile of the distribution. The estimate of the overall Gini index for the labour market is 0,31 and the estimates of Theil's L and T are 0,15 and 0,18 respectively. Looking at the inequality measures by groups of firms (columns ii-iv), all our indices suggest that there is more inequality in wages in the groups of firms we associate to higher productivity levels (exporters, multi-plant, and foreign-owned firms) than in the groups of firms we associate to lower levels of productivity. Yet, the decomposition of the Theil's T index, GE(1), into within and between components indicates that over 94% of overall inequality comes from dispersion in wages within groups of firms. That is, wages are relatively equally distributed across types of business strategy and almost all inequality comes from differences in wages between workers within the groups of firms.

[Table 4 about here]

Since inequality coming from wage differences within groups of firms has such a major contribution to overall inequality, compared to inequality coming from differences in wages between groups of firms, we will proceed with our analysis by decomposing the structure of the within inequality into factor sources.

3.2.3 Factor source decomposition of within inequality: the role of unobserved human capital and compensation policies

In this subsection we disaggregate within inequality into contributions of different factors. This type of decomposition analysis of inequality is important for understanding the main determinants of inequality. The determinants we consider are the set of components estimated from model (2), the proportional decomposition of wage inequality is computed as follows:

$$\frac{Cov(y, x\widehat{\beta})}{Var(y)} + \frac{Cov(y, \widehat{\theta})}{Var(y)} + \frac{Cov(y, \widehat{\psi})}{Var(y)} + \frac{Cov(y, \epsilon)}{Var(y)} = \frac{Var(y)}{Var(y)} = 1, \quad (7)$$

where $cov(\cdot)$ stands for covariance and $var(\cdot)$ stands for variance. The sum of the contributions of the observed time varying covariates ($x\widehat{\beta}$), the worker ($\widehat{\theta}$) and firm ($\widehat{\psi}$)

time-invariant effects, and the unexplained part of the model (ϵ) give the aggregate inequality value, and their proportional contributions add up to one (Shorrocks, 1982). Since we are interested in disentangling within inequality we perform this type of decomposition by groups of firms. The results are presented in Table 5. Columns (i) to (vi) report the decompositions by type high- low-roads. As mentioned in Section 2.2 we can decompose the time invariant effects of workers and firms into observable and unobservable components (see, e.g., equation (3) for the worker effects). Therefore, in Table 5 we perform such decompositions in order to identify the shares of inequality due to: (i) observed characteristics of workers and firms (including time-varying characteristics, and the time invariant gender, education, and industry affiliation of the firm) i.e., the share of observed covariates combines the effects of $\mathbf{x}\widehat{\beta}+\boldsymbol{\mu}\widehat{\eta}+\mathbf{k}\widehat{\kappa}$; (ii) total worker invariant effects $(\widehat{\theta})$ and unmeasured human capital $(\widehat{\alpha})$; total firm effects $(\widehat{\psi})$, and unmeasured compensation policies of the firm (firm effects net of industry effects: $\left(\widehat{\psi}-\widehat{\kappa}\right)$) this allows one to infer on the importance of education and gender on the contribution of worker time invariant effects, and the importance of industry affiliation on the contribution of firm time invariant effects.

[Table 5 about here]

Considering the economy overall (column i) wage differentials can be decomposed as follows: Unmeasured human capital $(\widehat{\alpha})$ accounts for the greatest share of the wage structure we observed in the economy, 42% of the difference in wages across workers is due to this component. The second most important factor explaining the differences in wages across workers are the observed characteristics of workers and firms, that account for 38% of the differentials. Within the observed component, education and gender account for about 17p.p. $(\widehat{\theta}-\widehat{\alpha})$ of the observed wage inequality and industry affiliation of the firm $(\mathbf{k}\widehat{\kappa})$ is responsible for about 7p.p. $(20.42-13.14)$ of the structure in wages. Compensation policies of the firm that differ from the compensation policies common to all firms within each industry $\left(\widehat{\psi}-\widehat{\kappa}\right)$ are responsible for 13% of the wage inequality. Finally, 6% of the estimated inequality is not explained either observed or unobserved components of the model. These results are in line with those obtained by Goux and Maurin (1999), AKM (1999) and Abowd, Finer and Kramarz (1999) in that they suggest that most of the differences in wages are accounted for unmeasured human capital.

As mentioned in the previous section, we have an interest in investigating what is behind the within inequality. That is, despite the partitions created according to different productivity proxies, we concluded that most inequality was generated within groups of the population rather than between groups. Hence, most of the observed wage differentials are not due to differences in productivity across firms but these exist within groups of high- and low-road industries and firms, they are transversal to the type of grouping we make. The decomposition of the differentials within groups of the population is presented in columns (ii) - (vi). Considering column ii, where we distinguish between the industries that pay the highest and lowest wage premia in the economy, we realize that within these two extremes observed characteristics of workers and firms are responsible for a smaller share of the wage differences (when compared to the economy overall, col. i), only 31% of the wage differentials are explained by this component. The loss in explanatory power of the observed characteristics is caused by the loss in importance of industry affiliation (probably because it was the factor used for the division). We can see that reduced explanatory power of industry affiliation if we compare the contribution of the firm effect $\left(\widehat{\psi}\right)$ and the net firm effect $\left(\widehat{\psi}-\widehat{\kappa}\right)$ which are essentially the same around 11% for the bottom 5, and 7% for the top 5. Within these groups of the population the largest share of the differential is accounted for by the worker time invariant characteristics ($\widehat{\theta}=66\%$), and more than 50% is due to unmeasured human capital ($\widehat{\alpha}$). When we split the population into deciles of firms, firm unobserved effects do not make a positive contribution to wage inequality (as expected, since this grouping is based on the distribution of estimated firm effects). Within the two deciles (p90 or p10) the firm compensation policies contribute to equalizing the differences in wages (contribution of -6%). On the other hand, time-invariant characteristics of the workers have their share increased and more 76% of the differentials in wages within these groups are accounted for worker time invariant effects, and at least 60% of that is accounted for unmeasured human capital. Therefore when we have very homogeneous groups of firms in terms of their compensation policies (across industries), worker abilities become the major factor underlying the wage structure. Our partitions of the population based on exporting, multiplant and foreign ownership as proxies for productivity are also informative in that we realize that in the more productive groups (exporters, multiplant and foreign owned) only about 5% of the wage structure is not considered by our model. The contribution of worker time invariant effects, in general, and of unmeasured human capital, in particular, is largest in the most productive groups than in their counterparts (and even larger than that observed in the economy overall). Although using cross-sectional data, (Doms et al. 1997) find a similar result in that they conclude that difference in time invariant unobserved worker quality are the main source of the cross-sectional correlation between productivity (technology use) and worker wages. For exporters and multi-establishment firms the contribution of observed characteristics of workers and firms is larger (38%) than that observed for the non-exporters and single-plant firms (35%). Firm compensation policies net of industry effects are also less powerful in determining wage differentials within these groups (11% for exporters and multiplant vs. 13% for the economy overall). In the case of foreign owned firms (column vi) there seems to be a trade-off in the importance of observed characteristics, which account for only 35% of the differentials (compared to 38% in the other partitions), and the firm compensation policies (net of industry effects) which account for almost 15% of the wage structure - the largest contribution across all types of partitions we have made.

Therefore, while in the previous section we have understood that most inequality comes from within productivity groups of firms, in this section we were able to demonstrate that most of the inequality observed within groups of firms (aggregated by productivity levels) is induced by unobserved differences in the workers' human capital. We have also realized that whilst industry affiliation explains about 7% of the differentials, firm-specific compensation policies are also important as they account for at least 11% of the wage structure. Therefore, although changing firms (or industry) may be a stepping stone for wage growth, investments in human capital may have greater effects for that matter.

4 Conclusion

Using Portuguese LEED *Quadros de Pessoal* data under the AKM framework, we estimate unobserved person- and firm-effects identifying high- and low-wage firms, and the industry wage structure. Then, we analyse differences and similarities of high-road and low-road industries and firms with respect to elements of human resources strategy. Our results indicate the existence of different wage policies across industries, after controlling for time-invariant unobserved characteristics of the workers, which suggests that the differentials estimated from cross-sectional data are not due to compositional effects of the workforce and that different firms pay different wages to workers with the same characteristics (measured or unmeasured). This means that substantial wage growth can be attained via mobility across industries. We also find considerable dispersion in firm compensation policies within both high- and low-road industries. Furthermore, the results suggest differences across the groups of firms both in terms the observed composition of the workforce and in terms of the estimated unobserved factors. We also conclude that time invariant unmeasured human capital is the major source of inequality within groups of firms, yet differences in compensation policies across firms (net of industry effects) are also a relevant factor. Therefore, it is by decreasing the differences in the workforce human capital (transversal skills, training, specific capital), that wage differences would be reduced.

References

Abowd, John M., Robert H. Creecy and Francis Kramarz (2002) "Computing person and firm effects using linked longitudinal employer-employee data." U.S. Census Bureau Technical Paper No. TP-2002-06.

Abowd, John M., Hampton Finer and Francis Kramarz (1999) "Individual and firm heterogeneity in compensation: an analysis of matched longitudinal employer-employee data for the State of Washington." in Haltiwanger, J. C., J. I. Lane, J. R. Spletzer, J. J. M. Theeuwes and K. R. Troske (editors): *The Creation and Analysis of Employer-Employee Matched Data.* North-Holland, Amsterdam: 3-24.

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

Audretsch, David B. and Talat Mahmood (1995) "New firm survival: new results using a hazard function.", Review of Economics and Statistics, 77(1): 97-103.

Bertrand, Marianne and Antoinette Schoar (2003) "Managing with style: the effect of managers on firm policies", Quarterly Journal of Economics, 118(4): 1169-1208.

Bloom, Nicholas and John Van Reenen (2007) "Measuring and explaining management practices across firms and countries", Quarterly Journal of Economics 122(4): 1351-1408.

Conyon, Martin J., S. Girma, S. Thompson and P. W. Wright (2002) "The productivity and wage effects of foreign acquisition in the United Kingdom.", *Journal of Industrial Economics,* 50(1): 85-102.

Cowell, Frank A. and C. V. Fiorio (2010) "Inequality decompositions: A Reconciliation", *Journal of Economic Inequality,* 9(4): 509-528.

Davidson, Carl; Fredrik Heynman, Steven Matusz, Fredrik Sjöholm and Susan Zhu (2010) "Globalization and imperfect labor market sorting.", Research Institute of Industrial Economics, IFN Working Paper No 856.

Davis, Steve J. and John Haltiwanger (1991) "Wage dispersion between and within U.S. manufacturing plants, 1963-1986", in Martin Baily and Clifford Winston (editors): *Brookings papers on Economic Activity, Microeconomics*:.115-200.

Doms, Mark; Timothy Dunne and Kenneth R. Troske (1997) "Workers, wages, and technology.", *Quarterly Journal of Economics,* 112(1): 253-290.

Ferreira, Priscila (2009) "The sources of interindustry wage differentials.", ISER WP N.2009-13, Colchester, University of Essex.

Fiorio, Carlo V. and Stephen P. Jenkins (2010) "INEQRBD: Stata module to calculate regression-based inequality decomposition.", Statistical Software Components S456960, Boston College Department of Economics.

Garicano, Luis and Paul Heaton (2007) "Information technology, organization, and productivity in the public sector: evidence from police departments.", *Journal of Labor Economics,* 28(1): 167-201.

Girma, Sourafel; Avid Greenaway, and Richard Kneller (2004) "Does Exporting Increase Productivity? A Microeconometric Analysis of Matched Firms.", *Review of International Economics,* 12(5): 855-866.

Gibbons, Robert and Lawrence Katz (1992) "Does unmeasured ability explain interindustry wage differentials?", *Review of Economic Studies, 59: 515-535.*

Goux, Dominique and Eric Maurin (1999) "Persistence of interindustry wage differentials: A reexamination using matched worker-firm panel data.", *Journal of Labor Economics*, 17(3): 492-533.

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

Hartog, Joop; Pedro T. Pereira and José C. Vieira (2000) "Inter-industry wage dispersion in Portugal.", *Empirica*, 27: 353-364.

Helpman, Elhanan; Oleg Itskhoki and Stephen Redding (2010) "Inequality and unemployment in a global economy", *Econometrica*, 78(4): 1239-1283.

Jenkins, Stephen P. (2010) "INEQDEC0: Stata module to calculate inequality indices with decomposition by subgroup.", Statistical Software Components S366007, Boston College Department of Economics.

Krueger, Alan B. and Lawrence H. Summers (1988) "Efficiency wages and the interindustry wage structure.", *Econometrica*, 56(2): 259-293.

Shorrocks, A. F. (1984) "Inequality decomposition by population subgroups.", *Econometrica*, 52(6): 1369-1385.

Syverson, Chad (2011) "What determines productivity?", *Journal of Economic Literature,* 49(2): 326-365.

Tor, Erikson (2012) "Progression of HR practices in Danish firms during two decades", unpublished paper, Aarhus University.

Woodcock, Simon (2008) "Wage differentials in the presence of unobserved worker, firm and match heterogeneity.", *Labour Economics*, 15(4): 772-794.

Appendix: Tables and Figures

Table 1: Summary statistics

Workers	Mean	Std Dev.	Macroeconomic	Mean	Std Dev.
ln(real hourly wage)	1.52	0.51	Region		
Age	37.74	11.10	North Coast	0.28	
Tenure	7.66	8.37	Center Coast	0.15	
Gender			Lisbon & Tagus Valley	0.42	
Male	0.57		Inland	0.07	
Female	0.43		Algarve	0.04	
Education			Islands	0.04	
ISCED1	0.46		Year		
ISCED2	0.22		2002	0.11	
ISCED3	0.21		2003	0.11	
ISCED5/6	0.11		2004	0.12	
Skill-level			2005	0.13	
Low	0.37		2006	0.13	
Medium	0.41		2007	0.13	
High	0.21		2008	0.14	
Type of contract of employment			2009	0.13	
Open-end	0.73				
Closed-end	0.27				
Type of work					
Full-time	0.97				
Part-time	0.03				
Firms					
ln(real sales volume)	15.22	2.60			
ln(firm size)	4.34	2.15			
Exporter					
No	0.53				
Yes	0.47				
Multi-establishment					
No	0.62				
Yes	0.38				
Ownership					
National	0.88				
Foreign	0.12				
No. of worker-year obs		13,385,663			
No. of workers		3,323,016			
No. of firms		246,564			

Table 2: Summary statistics - industries at SIC2 level

	Industry description	%
15	Manuf. of food, beverages and tobacco	3.69
17	Manuf. of textiles	2.91
18	Manuf. of wearing apparel; dressing and dying of fur	3.47
19	Tanning and dressing of leather; Manuf. of luggage, handbags, saddlery, harness and footwear	1.99
20	Manuf. of wood, products of wood & cork, except furniture; Manuf. of articles of straw & plaiting	1.43
21	Manuf. of pulp, paper and paper products	0.51
22	Publishing, printing and reproduction of recorded media	1.32
24	Manuf. of chemicals and chemical products, coke, refined petroleum products and nuclear fuel	0.97
25	Manuf. of rubber and plastic products	1.01
26	Manuf. of other non-metallic mineral products	2.12
27	Manuf. of basic metals	0.42
28	Manuf. of fabricated metal products, except machinery and equipment	2.96
29	Manuf. of machinery and equipment n.e.c	1.49
31	Manuf. of electrical machinery and apparatus n.e.c.	0.78
32	Manuf. of radio, television and communication equipment and apparatus	0.56
33	Manuf. of medical, precision and optical instruments, watches and clocks	0.22
34	Manuf. of motor vehicles, trailers and semi-trailers	1.32
35	Manuf. of other transport equipment	0.32
36	Manuf. of furniture; other manufacturing activities n.e.c.	1.72
37	Recycling	0.12
40	Production and distribution of electricity, gas, steam and hot water supply	0.25
41	Water collection, treatment and distribution	0.16
45	Construction	12.45
50	Sale. maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel	3.60
51	Wholesale trade and commission trade, except of motor vehicles and motorcycles	7.56
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	9.32
55	Hotels and restaurants	7.20
60	Land transport; transport via pipelines	3.12
61	Water transport	0.09
62	Air transport	0.47
63	Supporting and auxiliary transport activities; activities of travel agencies and other tourist activ.	1.46
64	Post and telecommunications	1.58
65	Financial intermediation, except insurance and pension funding	1.44
66	Insurance, pension funding and other complementary activities of social security	0.55
67	Activities auxiliary to financial intermediation	0.20
70	Real estate activities	0.67
71	Renting of machinery and equipment without operator, and of personal and household goods	0.31
72	Computer and related activities	0.98
73	Research and development	0.09
74	Other business activities	10.17
80	Education	1.74
85	Health and social work	4.64
90	Sewage and refuse disposal, sanitation and similar activities	0.21
91	Activities of membership organizations n.e.c.	0.72
92	Recreational. cultural and sporting activities	0.99
93	Other services	0.68

Table 3: High- and low-roads: descriptive statistics

	All	Industries		Firms		Firm: Exporter		Firm: Multi-plant		Firm: Ownership	
	(i)	(ii)		(iii)		(iv)		(v)		(vi)	
Observed covariates:		Bottom5	Top5	pct10	pct90	No	Yes	No	Yes	National	Foreign
ln(real hourly wage)	1.52	1.20	2.35	1.15	2.08	1.40	1.66	1.43	1.67	1.49	1.76
Woman	0.43	0.64	0.41	0.37	0.37	0.44	0.41	0.41	0.46	0.43	0.46
Age	37.74	37.58	39.01	38.06	38.86	37.6	37.92	37.85	37.58	37.97	36.05
Tenure	7.66	9.94	11.30	6.34	8.59	5.92	9.35	7.13	8.50	7.60	8.08
Close-end	0.27	0.18	0.10	0.24	0.22	0.31	0.23	0.28	0.27	0.28	0.26
ISCED											
2	0.22	0.16	0.13	0.22	0.18	0.21	0.22	0.21	0.23	0.21	0.24
3	0.21	0.08	0.45	0.19	0.31	0.19	0.24	0.17	0.28	0.20	0.30
56	0.11	0.02	0.34	0.07	0.24	0.10	0.12	0.09	0.13	0.10	0.16
Skill level											
medium	0.41	0.48	0.40	0.49	0.37	0.43	0.40	0.44	0.37	0.42	0.33
high	0.21	0.08	0.49	0.17	0.42	0.18	0.25	0.19	0.25	0.21	0.28
ln(real sales volume)	15.22	14.35	18.42	12.78	15.46	13.77	16.86	14.11	17.00	14.89	17.60
ln(size)	4.34	4.03	6.04	2.24	3.93	3.36	5.44	3.37	5.89	4.08	6.24
Exporter	0.47	0.64	0.54	0.17	0.47	–	–	0.36	0.64	0.42	0.86
Multi-plant	0.38	0.13	0.86	0.12	0.38	0.26	0.52	–	–	0.35	0.63
Foreign owner	0.12	0.08	0.21	0.02	0.18	0.03	0.22	0.07	0.20	–	–
Estimated effects:											
Observed covariates	-0.03	-0.05	0.10	-0.12	-0.01	-0.09	0.03	-0.08	0.03	-0.05	0.06
Worker effects (θ)	0	-0.21	0.44	0.19	0.11	-0.05	0.05	-0.04	0.06	-0.01	0.09
observed time invariant (η)	-0.06	-0.16	0.09	-0.07	0.03	-0.07	-0.04	-0.07	-0.04	-0.06	-0.02
unobs. human capital (α)	0.06	-0.05	0.35	0.27	0.08	0.02	0.09	0.03	0.10	0.05	0.11
Firm effects (ψ)	0.02	-0.08	0.28	-0.46	0.44	-0.00	0.04	0.01	0.03	0.01	0.07
Industry (κ)	0.02	-0.08	0.28	-0.01	0.09	0.01	0.03	0.01	0.04	0.02	0.03
Firm net of industry eff.	0	-0.00	0.00	-0.45	0.35	-0.01	0.01	0.00	-0.01	-0.01	0.04
Correlation of time invariant effects:											
θ vs. ψ	0.09	-0.15	-0.22	-0.58	-0.24	-0.05	0.23	-0.03	0.28	0.04	0.36
α vs. ψ	-0.01	-0.20	-0.24	-0.59	-0.28	-0.16	0.14	-0.14	0.19	-0.06	0.27
α vs. $(\psi - \kappa)$	-0.11	-0.20	-0.24	-0.60	-0.39	-0.23	0.02	-0.20	0.05	-0.16	0.17
No. of observations	13,385,663	1,442,767	338,963	366,645	1,063,719	7,111,232	6,274,431	8,256,113	5,129,550	11,770,622	1,615,041

Note: Since models were estimated with all variables (dependent and independent) in deviations from the grand mean E[Y] is zero in the model using the full sample, In practical terms, the expected value of wages is given by the sum of the constant (0.017) with the estimated effect of the observed covariates (-0.034) plus the estimated firm effect (0.017). Differences in the means of the estimated effects, across groups of the population, are statistically significant at 1% significance level.

Table 4: High- and low-roads: measures of inequality in real hourly wages

	All	SIC2	SIC2			Firms			Firm: Exporter		Firm: Multi-plant		Firm: Ownership	
	(i)	(ii)	(iii)			(iv)			(v)		(vi)		(vii)	
		(46 groups)	Bottom 5	middle 36	Top 5	P10	interdecile	P90	No	Yes	No	Yes	National	Foreign
Workers share (%)			11	87	2	2.74	89.31	7.95	53	47	62	38	88	12
Income share (%)			7	87	6	1.74	84.39	13.87	46	54	55	45	84	16
p90/p10	3.654		1.980	3.509	3.018	2.097	3.267	4.496	2.948	4.060	3.078	4.182	3.460	4.459
Gini	0.313		0.196	0.305	0.233	0.195	0.293	0.304	0.278	0.324	0.285	0.330	0.304	0.334
GE(0) Theil's L	0.155		0.072	0.146	0.090	0.067	0.137	0.151	0.126	0.164	0.131	0.170	0.146	0.171
GE(1) Theil's T	0.179		0.093	0.169	0.086	0.084	0.160	0.146	0.153	0.180	0.156	0.186	0.166	0.187
Within and between decomposition of Theil's T														
Within inequality GE(1)		0.135		0.159			0.157		0.168		0.170		0.173	
% of total inequality		75.4		89.1			87.7		93.9		95.0		96.7	
Between inequality GE(1)		0.044		0.019			0.022		0.011		0.009		0.005	
% of total inequality		24.6		10.9			12.3		6.1		5.0		2.8	

Note: The decompositions were computed using the Stata module INEQDECO by Jenkins (2010).

Table 5: High- and low-roads: contribution of estimated effects to within-group inequality (%)

	All	Industries		Firms		Firm: Exporter		Firm: Multi-plant		Firm: Ownership	
	(i)	(ii)		(iii)		(iv)		(v)		(vi)	
		Bottom5	Top5	pct10	pct90	No	Yes	No	Yes	National	Foreign
Observed ($\mathbf{x}\beta + \boldsymbol{\mu}\eta + \mathbf{k}\kappa$)	38.34	31.11	30.94	30.19	38.65	35.74	37.47	35.14	38.18	38.16	34.84
Time invariant effects											
Worker (θ)	59.90	67.27	65.82	78.77	76.38	57.44	64.42	59.23	62.42	59.84	64.55
Human capital (α)	42.37	49.59	54.40	63.75	59.59	39.45	46.32	40.98	45.15	42.48	45.53
Firm (ψ)	20.42	11.62	7.78	-2.38	0.97	23.18	18.55	21.82	20.04	20.05	22.02
Firm net of industry effects ($\psi - \kappa$)	13.14	11.04	7.07	-5.82	-6.04	16.67	10.89	16.64	11.09	12.83	14.45
Residual	6.15	8.26	7.59	11.88	7.80	8.14	5.32	7.24	5.58	6.53	5.18

Note: The observed component includes the contribution of the observed time varying covariates and the time invariant observed covariates (gender, education and industry). The contributions of the observed covariates, the unmeasured human capital, the firm effect net of industry effects and the residual add up to 100. These regression based decompositions were computed using the Stata module INEQRBD by Fiorio and Jenkins (2010).

Figure 1: Raw interindustry wage differentials, cross sectional specification

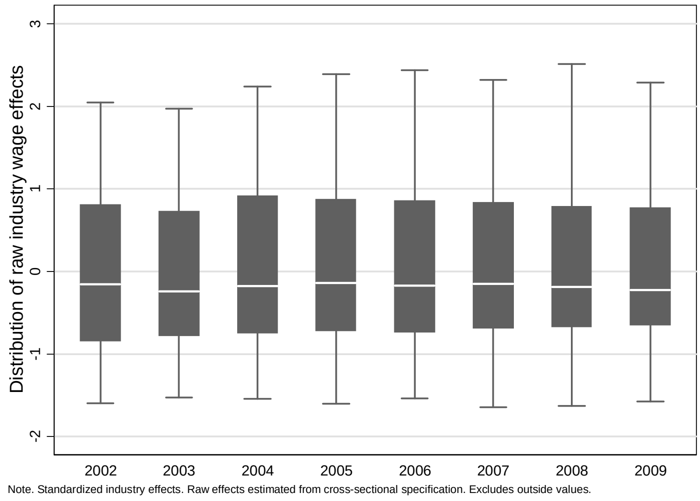

Figure 2: "True" interindustry wage differentials, AKM specification

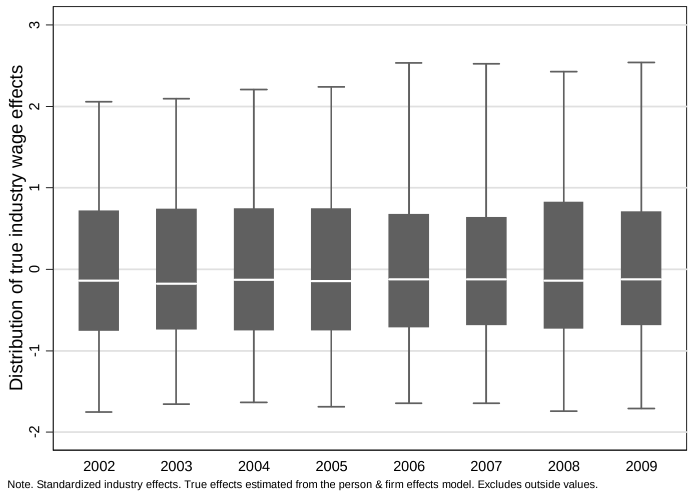

Figure 3: Association between the dispersion of firm effects and the industry wage premia