Available online at www.sciencedirect.com
SCIENCE DIRECT®
Economics Letters 83 (2004) 157–163
economics letters
www.elsevier.com/locate/econbase

Industry wage premia: evidence from the wage distribution

Pedro S. Martins*

Department of Economics, University of St. Andrews, St. Andrews, Fife KY16 9AJ, UK

Received 12 November 2001; received in revised form 23 October 2003; accepted 3 November 2003

Abstract

We argue that, under the unobserved quality explanation of industry wage differentials, industries with high average premia will have even higher premia among high-wage workers. We find, however, that the OLS premia correctly illustrate the industry impact across the distribution.

Keywords: Inter-industry wage differentials; Quantile regression; Unobserved heterogeneity; Portugal

JEL classification: J31; C29

1. Introduction

A clear stylised fact in labour economics concerns the inter-industry wage structure. Both across time and across countries, some specific industries tend to pay their workers more than other industries, even after controlling for a large set of workers' characteristics (see Dickens and Katz, 1987). Taken at face value, these findings clearly contradict the law of one price, thus suggesting that labour markets are not appropriately described by a competitive framework.

Such puzzle has sparked a non-competitive approach to labour markets, in general associated with the concept of efficiency wages (see Krueger and Summers, 1988). However, more sophisticated competitive explanations, involving either compensating differentials or unobserved quality differences across workers, have also been proposed.

This paper focuses on the latter explanation, which argues that high-wage industries are over-represented with high-ability workers. According to this view, the standard results on inter-industry wage

* Tel.: +44-1334-46-2423; fax: +44-1334-46-2444.
E-mail address: pm31@st-andrews.ac.uk (P.S. Martins).

0165-1765/$ - see front matter
doi:10.1016/j.econlet.2003.11.002

differentials are biased due to lack of controls for a number of workers' characteristics that are relevant in the wage determination process (e.g. ability, motivation, industry-specific skills).[1]

A possible implication of this competitive framework, which is addressed in this paper, is that the role of industry affiliation upon wages should differ depending on the specific part of the conditional distribution of wages being addressed. In particular, in the top of the conditional wage distributions, where the high-skilled workers can be found, the wage impact of high-wage industries should be particularly noticeable. Conversely, this wage premium should be substantially lower or even non-existent at the bottom of the wage distribution.

Under this approach, the traditional OLS results would be driven by a longer upper tail of the conditional wage distribution for high-wage industries. This longer upper tail would be due to the over-representation of high-skilled workers in those industries, as the unobserved quality explanation assumes. An alternative case, more difficult to rationalise with unobserved heterogeneity explanations, is that the conditional wage distribution is subject to pure location shifts, where the OLS results would be similar to the industry impacts across the wage distribution. This is in line with, for instance, the predictions of the 'fairness' models, where employers consider profitable to drive up the wages of all workers in order to promote cohesion within their firms (see Akerlof, 1984).

This hypothesis—that unobserved heterogeneity is behind industry wage differentials—is tested here by drawing on the results of wage equations estimated by quantile regression (Koenker and Bassett, 1978). Rather than fitting the equation through the mean of the dependent variable, quantile regression considers the impact of the regressors at specific quantiles of the distribution of the dependent variable. This method is insightful whenever conditional distributions of the dependent variable are not simply characterised by their mean.

Our approach involves the regression of an extended wage equation, controlling for a large number of both demand- and supply-side variables, at the mean, median and at the 'top' and 'bottom' quantiles (10th and 90th percentiles) of the distribution of wages. We then test the unobserved worker quality hypothesis by studying the difference in the returns across the distribution for high- and low-wage industries. We also examine this hypothesis by examining the correlation between the industry premia at those quantiles and by considering the correlation between the mean industry premia and the difference between the industry premia at the top and bottom quantiles.[2]

Given the discussion above, only a bigger difference between the top and the bottom of returns for high-wage industries than low-wage industries and a highly positive correlation between the mean and top quantile premia would be consistent with the unobserved worker quality explanation.[3] This would signify that the high average premia associated with high-wage industries are driven by particularly high industry returns at the top of the conditional wage distributions, precisely where the highest unobserved ability workers would be expected to be found.

[1] This matter is far from being settled. See the conflicting evidence in Gibbons and Katz (1992); Abowd et al. (1999), inter alia.

[2] Kahn (1998) also draws upon quantile regression results in the inter-industry wage differentials context. However, his interest lies on comparing the dispersion of industry premia along the wage distribution and across countries. Here we focus on the correlation of industry premia at different points of the wage distribution.

[3] An equivalent test involves the correlation between, on the one hand, the mean premia and, on the other hand, the difference between the top and bottom quantiles. This correlation would be high under the unobserved ability hypothesis.

In this paper, we find, using Portuguese data presented in Section 2, however, strong correlations between the mean premia and both the top and bottom premia. This result, presented in Section 3, therefore, argues against the view that the commonly-documented average industry premia are simply a consequence of unobserved heterogeneity.

2. Data

We draw upon a sample from a Portuguese matched employer-employee data set, 'Quadros de Pessoal' (Personnel Records), from 1995. This includes a large amount of information on a representative sample of 41,058 employees, covering both traditional human capital variables (schooling, experience and tenure) and also demand-side evidence (such as occupation level or firm size).[4]

Also importantly, this data source is particularly reliable, as the information is subject to scrutiny by the Portuguese Ministry of Employment, given the latter's regulation duties. Moreover, since the required information is filled in by the employer, the workers' industry affiliation variable is likely to be much more robust to measurement error problems that affect other studies.

It should also be mentioned that Portugal is a country particularly suitable for the purposes of this paper, given its large inter-industry wage differentials, as carefully documented in Hartog et al. (2000). This result, from a competitive approach, would suggest a considerable scope for unobserved heterogeneity, thus strengthening the case for not rejecting the hypothesis considered in this paper.

Table 1 presents some descriptive statistics on the sample used in this paper. The full set of control variables includes: dummies for different educational levels (nine levels), tenure and experience (and their squares), a female dummy, log monthly hours worked, 43 two-digit sector dummies, 26 two-digit job types, eight hierarchical job classification dummies, log firm size (number of workers), six firm region dummies and three dummies for firm ownership (private, public or foreign). The dependent variable is the log of total gross monthly wages.[5]

3. Results

In this section, we present our findings concerning possible differences between average industry wage premia and such premia at the top and bottom of the wage distribution. As outlined in Section 1, our intuition is that an unobserved quality interpretation of industry differentials would be associated with higher premia at the top than at the bottom of the conditional wage distribution. This difference would follow from an over-representation of high unobserved quality workers at the top of the high-wage industry conditional wage distributions.

Table 1 also presents the average and median industry wage coefficients, in which the industries are ranked in increasing order of the OLS premia. As predicted, we find a considerable amount of dispersion across industries, even though we control for a particularly large set of variables. For example, OLS

[4] See Hartog et al. (2000), inter alia, for a more detailed description of the data set.

[5] Full descriptive statistics, as well as the entire results from the regressions presented in Section 2, can be obtained upon request.

Table 1
Industry shares, and OLS and median results

Industry	Sample share (%)	OLS		50th percentile	
		Coefficient	S.E.	Coefficient	S.E.
Textiles	5.7	− 0.1524	0.011	− 0.1374	0.0101
Clothing	6.3	− 0.142	0.0116	− 0.1191	0.0107
Furniture	2.0	− 0.1323	0.0145	− 0.1197	0.0133
Hotels and restaurants	6.1	− 0.0858	0.0119	− 0.0801	0.011
Other services	0.6	− 0.0626	0.0241	− 0.0523	0.0222
Leather and related products	3.4	− 0.0513	0.0133	− 0.0376	0.0123
Retail (except cars)	8.1	− 0.0466	0.0103	− 0.0339	0.0095
Machines and electric devices	1.4	− 0.0399	0.0182	0.0177	0.0168
Other services to firms	4.5	− 0.0352	0.0114	− 0.0268	0.0105
Food and drink	4.5	− 0.0252	0.0108	− 0.0139	0.01
Machines and equipment rental	0.2	− 0.0232	0.0373	− 0.0111	0.0343
Timber and cork	1.9	− 0.0141	0.0141	− 0.0015	0.013
Metal products	3.1	− 0.0125	0.0124	0.0125	0.0115
Health and social action	2.4	− 0.007	0.0157	0.029	0.0145
Building	10.1				
Land transport and pipelines	3.2	0.0028	0.0132	0.014	0.0122
Metallurgic industries	0.5	0.0121	0.0254	0.0377	0.0234
Car retail and repairing	3.8	0.0135	0.0113	0.0305	0.0104
Machines and related equipment	1.8	0.0348	0.0147	0.0739	0.0136
Medical devices	0.3	0.0364	0.0328	0.0384	0.0301
Radio, TV and communication equipment	0.7	0.0499	0.0224	0.1047	0.0207
Education	1.6	0.0613	0.0182	0.0988	0.0168
Gross retail	7.7	0.0655	0.0094	0.0755	0.0087
Cars	0.9	0.0685	0.0195	0.0728	0.018
Rubber and plastic	1.0	0.0813	0.0187	0.0812	0.0173
Cultural, sport and leisure activities	0.7	0.0825	0.0216	0.1043	0.0199
Other transport material	0.7	0.0836	0.0218	0.1462	0.0201
Non-metal mineral products	2.9	0.0875	0.0128	0.1006	0.0118
Post offices and telecommunications	2.0	0.09	0.018	0.1374	0.0166
Various social activities	0.4	0.0927	0.0284	0.1603	0.0262
Other mining	0.6	0.104	0.0217	0.1743	0.0199
Property-related activities	0.4	0.1296	0.0288	0.1235	0.0265
Computer and related activities	0.2	0.1458	0.0377	0.2099	0.0346
Editing and printing	1.5	0.1566	0.0168	0.1224	0.0155
Paper and cardboard	0.7	0.1577	0.0215	0.1472	0.0198
Chemical products	1.3	0.169	0.0163	0.1937	0.015
Coal and oil	0.2	0.1901	0.0914	0.2246	0.0815
Air transport	0.5	0.222	0.0279	0.1898	0.0257
Other financial services	0.1	0.2721	0.0456	0.3747	0.0418
Financial services	3.6	0.2862	0.024	0.3287	0.0221
Travel agencies and related activities	1.0	0.3169	0.0178	0.2971	0.0164
Electricity, water and gas	0.9	0.3663	0.0232	0.4353	0.0214
Insurance and pension funds	0.7	0.3935	0.0207	0.4555	0.0191
R^2		0.666			
F-statistic		825			

Extra controls considered but not reported here (see main text).

Table 2
Industry premia along the distribution and their difference

Industry	10th percentile		90th percentile		Difference	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	*P*-value
Textiles	− 0.1452	0.0107	− 0.2084	0.0212	− 0.0632	0.036
Clothing	− 0.127	0.0117	− 0.1998	0.0224	− 0.0728	0.004
Furniture	− 0.1127	0.0142	− 0.1945	0.0276	− 0.0818	0.019
Hotels and restaurants	− 0.0871	0.0117	− 0.0961	0.0239	− 0.009	0.830
Other services	− 0.0551	0.0238	− 0.0898	0.0476	− 0.0347	0.632
Leather and related products	− 0.0192	0.0137	− 0.1183	0.026	− 0.0991	0.002
Retail (except cars)	− 0.0716	0.0106	− 0.0528	0.0196	0.0188	0.611
Machines and electric devices	0.0344	0.0179	− 0.1274	0.0362	− 0.1618	0.000
Other services to firms	− 0.0167	0.0115	− 0.0219	0.0229	− 0.0052	0.885
Food and drink	− 0.0653	0.0107	− 0.0309	0.0207	0.0344	0.171
Machines and equipment rental	− 0.0279	0.0356	− 0.013	0.07	0.0149	0.890
Timber and cork	− 0.0512	0.0137	− 0.0281	0.0269	0.0231	0.438
Metal products	0.0082	0.0121	− 0.0377	0.0241	− 0.0459	0.128
Health and social action	0.0478	0.0154	− 0.0674	0.0321	− 0.1152	0.029
Land transport and pipelines	0.03	0.0124	− 0.0303	0.0276	− 0.0603	0.141
Metallurgic industries	0.118	0.0246	− 0.0982	0.049	− 0.2162	0.000
Car retail and repairing	− 0.0079	0.0112	− 0.0337	0.0215	− 0.0258	0.325
Machines and related equipment	0.0555	0.0145	− 0.0162	0.0282	− 0.0717	0.087
Medical devices	0.0634	0.032	− 0.0762	0.0625	− 0.1396	0.240
Radio, TV and communication equipment	0.1697	0.0223	0.0011	0.0438	− 0.1686	0.006
Education	0.0518	0.0187	− 0.0262	0.0343	− 0.078	0.025
Gross retail	0.0305	0.0095	0.0535	0.0175	0.023	0.438
Cars	0.0699	0.0193	0.1277	0.0371	0.0578	0.173
Rubber and plastic	0.1037	0.0182	0.0266	0.0364	− 0.0771	0.052
Cultural, sport and leisure activities	0.0596	0.0213	0.0439	0.0411	− 0.0157	0.702
Other transport material	0.2296	0.0218	− 0.0598	0.0416	− 0.2894	0.000
Non-metal mineral products	0.0825	0.0132	0.0643	0.0248	− 0.0182	0.490
Post offices and telecommunications	0.2175	0.0179	0.0162	0.0361	− 0.2013	0.000
Various social activities	0.1679	0.0273	0.0165	0.0543	− 0.1514	0.025
Other mining	− 0.0162	0.0214	0.1182	0.0406	0.1344	0.013
Property-related activities	0.037	0.0279	0.2254	0.0547	0.1884	0.011
Computer and related activities	0.0178	0.0369	0.1126	0.0728	0.0948	0.532
Editing and printing	0.0773	0.0165	0.1844	0.0326	0.1071	0.006
Paper and cardboard	0.0543	0.0214	0.1823	0.0398	0.128	0.286
Chemical products	0.1602	0.0162	0.1213	0.0314	− 0.0389	0.544
Coal and oil	0.3926	0.0786	− 0.1283	0.1525	− 0.5209	0.000
Air transport	0.1275	0.0267	0.3094	0.0544	0.1819	0.017
Other financial services	0.1622	0.0425	0.1648	0.0832	0.0026	0.984
Financial services, except insurance and pension funds	0.528	0.0265	0.1277	0.0475	− 0.4003	0.000
Travel agencies and travel related activities	0.3027	0.0175	0.3564	0.0345	0.0537	0.222
Electricity, water and gas	0.5113	0.0221	0.2202	0.0505	− 0.2911	0.000
Insurance and pension funds	0.6371	0.0198	0.1812	0.0401	− 0.4559	0.000

Extra controls considered but not reported here (see main text).

Table 3
Correlation between OLS and QR coefficients

	Correlation coefficient	Spearman coefficient
A, 10th percentile	0.861	0.838
B, 50th percentile	0.98	0.98
C, 90th percentile	0.834	0.82
Diff (= C − A)	− 0.239	0.025

wage differentials reach a maximum difference of 73% (when the textiles and the insurance and pension funds industries are compared).

In Table 2, we present the industry premia at the 10th and 90th percentiles and the difference between the two percentiles at each industry. The standard errors (S.E.) of the latter are obtained via a bootstrapped variance-covariance matrix that includes between quantiles blocks. We find that, in most cases, the returns are lower at the top of the distribution and their difference is not significant.

Moreover, the difference in the returns between the top and the bottom of the distribution is not bigger for higher-paying industries than for their lower-paying counterparts: whereas the average difference is − 0.043 for the 15 lower-wage industries (defined in terms of the OLS results), the same average is − 0.069 for the 15 higher-wage industries.[6] According to our "unobserved heterogeneity" hypothesis, we would expect, however, that the average differential would be *more positive* for the group of industries with higher average premia.

Table 3 presents standard and Spearman correlations among the industry coefficients from the different estimation methods and also the difference between the coefficients from the top and bottom quantiles. Again, and unlike what was hypothesised to be the case under the unobserved quality approach, it is found that OLS results are strongly correlated with those at the two extreme points of the wage distribution (the 10th and 90th percentiles) and also the 50th percentile. In all cases, correlation coefficients are above 0.80. Moreover, as suggested by a previous analysis, the correlation between the OLS coefficients and the differences between returns at the top and bottom of the wage distribution is very low (and negative).

4. Conclusion

We present a contribution to the inter-industry wage differentials debate. In particular, we suggest a test of the unobserved quality explanation, drawing on quantile regression. To the extent that high-wage industries draw disproportionately more on high-ability workers, as implied by the above hypothesis, the industry wage premia at the top of the conditional wage distributions should be higher than those at the bottom of the same conditional wage distributions.

We test this implication by drawing on Portuguese data, a country characterised by particularly high industry wage differentials. We find that quantile regression estimates of industry effects, at the bottom, middle or top of the conditional wage distribution, strongly resemble those obtained by using OLS. Our results, therefore, suggest that unobservable differences across workers are not a critical element in

[6] We thank the referee for suggesting this comparison.

explaining industry wage premia. Non-competitive forces may thus play an important role in the wage determination process.

Acknowledgements

I thank, without implicating, Ian Walker and Reamonn Lydon and one anonymous referee for their critical comments. Financial support from *Fundação para a Ciência e a Tecnologia* (SFRH/BD/934/2000 and POCTI/ECO/33089/99) and the British Council and logistical support from *Banco de Portugal* are gratefully acknowledged.

References

Abowd, J., Kramarz, F., Margolis, D., 1999. High-wage workers and high-wage firms. Econometrica 67, 251–333.

Akerlof, G., 1984. Gift exchange and efficiency–wage theory. Four views. American Economic Review 74, 79–83.

Dickens, W.T., Katz, L.F., 1987. Inter-industry wage differences and industry characteristics. In: Lang, K., Leonard, J. (Eds.), Unemployment and the Structure of Labor Markets. Basil Blackwell, New York, pp. 48–89.

Gibbons, R., Katz, L., 1992. Does unmeasured ability explain inter-industry wage differentials? Review of Economic Studies 59, 515–535.

Hartog, J., Pereira, P.T., Vieira, J.C., 2000. Inter-industry wage dispersion in Portugal: high but falling. Empirica 27, 353–364.

Kahn, L., 1998. Collective bargaining and the interindustry wage structure: international evidence. Economica 65, 507–534.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.

Krueger, A.B., Summers, L.H., 1988. Efficiency wages and the inter-industry wage structure. Econometrica 56, 259–293.