Working Papers, Centre for Globalisation Research

No. 16 (July 2008)

ISSN 1755-5892

Is firm performance driven by fairness or tournaments? Evidence from Brazilian matched data

by

Luiz. A. Esteves and Pedro S. Martins

Is firm performance driven by fairness or tournaments? Evidence from Brazilian matched data*

Luiz A. Esteves†

Universidade Federal do Paraná & Universitá di Siena

Pedro S. Martins‡

Queen Mary, University of London & CEG-IST, Lisbon & IZA, Bonn

July 2, 2008

Abstract

Theory and evidence are ambiguous about the effect of within-firm wage inequality on firm performance. This paper tests empirically this relationship drawing on detailed Brazilian matched employer-employee panel data, considering alternative measures of inequality and performance and different estimation methods. We find overwhelming evidence of a positive relationship between wage dispersion and firm performance when using cross-section analysis, especially in manufacturing. However, this relationship is weakened when controlling for firm time-invariant heterogeneity.

Keywords: Tournaments, Incentives, Equity, Wage Dispersion.

JEL Codes: D31, J31, J33, J41, J53.

---

*We are grateful for the feedback from Eduardo P. Ribeiro and participants in a workshop at IPEA, Brasilia. All remaining errors are ours.

†Email: esteves@ufpr.br

‡Corresponding author. Email: p.martins@qmul.ac.uk. Address: School of Business and Management, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom. Phone: +44/0 2078827472. Fax: +44/0 2078823615. Web: http://webspace.qmul.ac.uk/pmartins.

1 Introduction

Firms face important choices when designing incentives for workers, namely in terms of pay structures. To the extent that (perceptions of) 'fairness' is an important issue from the workers' point of view, large differences in pay levels for workers of similar characteristics may be detrimental for firm performance (Akerloff & Yellen 1988, 1990). On the other hand, workers' effort may be particularly sensitive to the potential of earning higher wages - as in 'tournaments' (Lazear & Rosen 1981) - in which case the relationship between pay inequality and firm performance will be positive.

This question, also posed in the title of our paper, has been examined empirically in several occasions with contrasting results. On the one hand, several studies present evidence supporting the 'tournaments' views (Main et al. 1993, Winter-Ebmer & Zweimuller 1999, Eriksson 1999, Hibbs & Locking 2000, Lallemand et al. 2004, Heyman 2005, Lallemand et al. 2007). On the other hand, there are also some studies supporting the 'fairness' view , although their number is smaller (DeBrock et al. 2001, Martins forthcoming).[1]

Our contribution to this empirical literature is two-fold. First, we draw on detailed matched employer-employee panel data. The potential gain in insight from this type of data sets is obvious, as one can build much better measures of worker and firm heterogeneity. As far as we know, only Martins (forthcoming) has used similar data sets before in this context; and few papers consider conditional measures of inequality, in which worker characteristics are partialed out from wages (Winter-Ebmer & Zweimuller 1999, Heyman 2005, Lallemand et al. 2007, Martins forthcoming). Second, we provide evidence for developing countries, which so far have been completely overlooked in this literature. However, in many respects the analysis of these countries - Brazil, in our case - is of particular interest, as the institutional settings there may allow for greater flexibility in human resource management, including pay setting, and therefore facilitate experimentation in terms of different HRM approaches.

In our analysis, we match three different data sets (RAIS, a matched employer-employee panel data set; PIA, a survey of manufacturing sector firms; and PAS, a similar survey for service sector firms), over the period 1998-2001. As far as we know, this is one of the first papers that considers both the manufacturing and the services sectors. We also consider alternative measures of within-firm inequality and of firm performance (profits and added

[1] Leonard (1990) does not establish any significant relationship. See Lallemand et al. (2007) also for a very detailed survey of the literature.

value per worker). Moreover, for the benefit of robustness, we employ a large set of different econometric models: OLS, quantile regression, instrumental variables, fixed effects and both fixed effects and instrumental variables.

The remainder of this paper is organized as follows. Section 2 presents the data, some descriptive statistics and the measures of inequality used. Sections 3 and 4 present the results. Finally, Section 5 concludes.

2 Data

As mentioned above, we use the following data sets: RAIS, a matched employer-employee panel data set; PIA, a survey of manufacturing sector firms; and PAS, a similar survey for service sector firms. RAIS ('Relação Anual de Informações Sociais', Annual Social Data Report), is a census of all firms and all their formal-sector employees in Brazil conducted by the Ministry of Labour. The data include detailed information about each employee (wages, hours worked, education, age, tenure, gender, worker nationality, etc) and each firm (industry, region, size, establishment type, etc) in each year, plus a unique identifier for each employee, each establishment and each firm.[2]

PIA ('Pesquisa Industrial Anual', Annual Manufacturing Survey) and PAS ('Pesquisa Anual dos Servi cos', Annual Services Survey) are yearly establishment surveys covering the entire country, conducted by IBGE (Brazilian Statistics Agency). They cover all establishments (census) with 30 or more workers (20 or more in the case of PAS), plus a random sample of establishments whose size ranges from 5 to 29 workers (5 to 19 in the case of PAS). Moreover, the surveys collect information on labour inputs, labour costs, turnover, production level and other variables.

PAS is ongoing since 1998 while PIA, at least in its current format, is ongoing since 1996. Here, we examine the 1998-2001 period[3] and consider a sample comprising a balanced panel of 7,689 firms, of which 4,990 are in the manufacturing sector and 2,699 in the services sector. The sample was constructed so that every firm selected has data for at least 50 employees in each year. As will be explained later, this restriction is imposed to allow sufficient degrees of freedom in the estimation of firm-year wage regressions. Moreover, in order to address outliers, we also restrict the sample to firms whose profits lie on the central 95% of the profits

[2] See Appendix D for more details on RAIS. See also Martins & Esteves (2006) for a recent application.
[3] All nominal variables are deflated by the Brazilian consumer price index (INPC).

per worker distribution every year throughout the study period.

The RAIS data is used twice: first, worker-level information is drawn to compute a conditional measure of inequality based on the residuals of wage regressions; second, the means of the human capital variables per firm are used in firm performance regressions. The first stage above involves individual information for nearly 1,300,000 workers-year (approximately 740,000 from the manufacturing sector and approximately 560,000 from the service sector). The list of variables used in the second stage of regressions, i.e., in the firm performance regressions, is shown in Table 9. Tables 1, 10 and 11 show the descriptive statistics of these variables for the whole set of firms, for the set of manufacturing firms and for the set of service firms, respectively.[4]

2.1 Inequality measures

We use two main types of measures of wage dispersion, namely, conditional and unconditional. Unconditional measures, which do not contemplate worker heterogeneity in terms of the observed characteristics of human capital (educational level, gender, job tenure, experience, and age), are considered in three different ways: the standard deviation of wages of firm j in year t; the coefficient of variation of firm j in year t; and the ratio between the minimum and maximum wage observed at firm j in year t.[5]

In order to take into account the extent of wage dispersion across homogeneous workers (the conditional measure), we use the standard error of a wage regression (σ), as in Winter-Ebmer & Zweimuller (1999). In this case, the analysis is performed in two stages, the first consisting in estimating a wage regression for each firm-year, according to the following simple specification:

$$ln\ w_i = X_i'\beta + \epsilon_i, \tag{1}$$

where lnw_i is the logarithm of the hourly wage of individual i; and X_i is a vector of variables capturing at least some of the attributes of worker i (educational level, gender, job tenure and

[4]Some findings follow from the analysis of Tables 1, 10 and 11: service firms employ on average more workers than do manufacturing firms; except for 1998, the profits per worker in manufacturing firms were higher than those in service firms; profit dispersion in the manufacturing sector was larger than in the service sector; manufacturing firms were more unequal than those in the service sector; most measures of inequality - described below - increased considerably over the period; manufacturing and services present similar means for education and female participation; workers in the manufacturing sector are younger, but their tenure is longer.

[5]An obvious limitation of unconditional measures of inequality is that worker heterogeneity in a given firm will imply wage dispersion that is not related to any incentive policy.

age). After estimating the regression for each firm-year, the conditional measure of inequality is the standard error of wage regression (σ), as specified below:

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{n} \left[ln\ w_i - \widehat{ln\ w_i}\right]^2}{(n-k)}}, \tag{2}$$

where σ_j represents the standard error of wage regression for firm j; $\widehat{ln\ w_i}$ is the estimated value of the logarithm of the hourly wage of individual i; n is the number of observations in the firm (and year); and k is the number of parameters to be estimated.

The second stage considers the role of wage dispersion variables (either conditional or unconditional) as firm performance determinants. Unlike in wage regressions of the first stage (at the worker level), performance regressions include information about the human capital of the workforce aggregated at the firm level, i.e., the means of human capital variables per firm.

3 Cross-Sectional Analysis

This section describes the results of performance regressions. As mentioned in the introduction, we consider here three different estimations methods: OLS, quantile regression and instrumental variables.

3.1 OLS

The econometric model follows the functional specification in Lallemand et al. (2004) and is as follows:

$$ln(P/n)_j = \beta M_{zj} + X_j'\delta + F_j'\varphi + \epsilon_j, \tag{3}$$

where the dependent variable is the logarithm of profit (or added value) per worker at firm j; M_{zj} is the measure of inequality or wage dispersion z of firm j, in which z corresponds to either sigma, the standard deviation, the coefficient of variation or the max-min ratio; X is the vector of variables that capture characteristics of the workers employed at firm j; and F is the vector of variables that capture characteristics of firm j.

Cross-sectional regressions were obtained for each year and for each one of the four wage dispersion measures, thus totaling 16 different regressions. The results are presented in Table

2, where we find that all coefficients of the wage dispersion measures have a positive sign. Moreover, only in the cases of the sigma and max-min ratio in 1998, or for the sigma, coefficient of variation and maxmin ratio in 1999 are the results not significant.

However, one issue with the functional specification above is that when using the logarithm of profits per worker as the dependent variable, the sample is restricted to firms with positive profits. This implies a loss of a large number of annual observations. An alternative approach is the use of the logarithm of value added as the dependent variable (Lallemand et al. 2007).

We follow this alternative approach and obtain cross-sectional analysis for each year and for each one of the four wage dispersion measures, using now as dependent variable the logarithm of the added value per worker. Similarly to the previous analysis, all 16 coefficients have a positive sign, but now all are statistically significant at the 1% level - see Table 3.[6]

Tables 4 and 5 show the results of the cross-sectional regressions separately for the manufacturing and services sectors, respectively. All coefficients have a positive sign and are significant at the 1% significance level. Except for the coefficient of the maxmin variable for 1998, all other coefficients are higher for the manufacturing sector than for the service sector.

Overall, the evidence above suggests a positive relationship between within-firm wage dispersion and economic performance for Brazilian firms. It also suggests that incentives generated by wage inequality are more effective in the case of manufacturing firms.

The second piece of evidence seems to corroborate the argument by Milgrom (1988) and Milgrom & Roberts (1990). The idea is that white-collar workers (more common work in the services sector) may withhold information from managers in order to increase their influence (e.g. to get a promotion) and to engage in rent seeking activities instead of productive work. A related view (Lazear 1989, 1995) is that service firms have a larger number of non-cooperative workers ('hawks') compared to the manufacturing sector, who would have a larger number of cooperative workers ('doves'). In both cases, some wage compression will be desirable for these workers.

The results presented so far indicate strong evidence in favor of the hypothesis of a positive relationship between within-firm wage dispersion and economic performance for the Brazilian case. This evidence is corroborated by unconditional measures and also by the sigma condi-

[6] By comparing the results in Tables 2 and 3, the specification with the logarithm of the added value (Table 3) improves the statistical significance of the tests and also prevents the large loss of observations comparatively to the results in Table 2 (approximately 2,400 observations of firms per year). The next subsections use only the specifications with the logarithm of the added value per worker as dependent variable.

tional measure of inequality.[7] However, note that the results obtained from OLS estimators may be biased. The next sections will deal with these problems by considering alternative econometric methods.

3.2 Quantile Regression

OLS estimates are based on the mean conditional distribution of the dependent variable. This approach implicitly assumes that the possible differences in the impact of exogenous variables across the conditional distribution of the dependent variable are negligible. In other words, if the exogenous variables exert some influence on other parameters of the conditional distribution of the dependent variable than the mean, an OLS approach will be misleading.

The aim of this section is to assess the relationship between wage inequality and firm performance using quantile regression (Koenker & Bassett (1978); see Martins & Pereira (2004) for an application. Differently from OLS, quantile regression characterizes the conditional distribution of the dependent variable, rather than assuming that the impact of the exogenous variables is the same along the distribution of the dependent variable. We believe this analysis can be particularly insightful in our context as it will highlight differences across firms of different efficiency levels.

In our context of firm performance equations, the quantile regression model can be defined as:

$$ln(P/n)_{\theta j} = Z_j'\beta_\theta + u_{\theta j}, \tag{4}$$

$$Quant_\theta[ln(P/n)_j|X_j] = X_j'\beta_\theta, \tag{5}$$

where $\mathbf{Z}_j$ is a vector of exogenous variables (which include M_j, X_j and F_j) and β_θ is a vector of parameters. $Quant_\theta[ln(P/n)_j|\mathbf{Z}_j]$ denotes the θth conditional quantile of $ln(P/n)$ given Z. In this paper, we present the coefficients obtained for the 10th, 25th, 50th, 75th and 90th percentiles (dubbed P10, P25, P50, P75 and P90, respectively).

Tables 6, 12 and 13 show the results of quantile regressions for all firms, for manufacturing firms and for firms in the service sector, respectively. All coefficients have a positive sign and, in most cases, are statistically significant. Some cases of insignificant coefficients were found in regressions concerning the 10th percentile of the conditional distribution (P10).

[7] As previously mentioned, the sigma measure is preferable to any other measure as far as the purpose of this study is concerned. The next sections will only consider specifications based on sigma.

Overall, the quantile regression results suggest that the coefficients of the relationship between wage inequality and firm performance increase with conditional performance. This finding strengthens earlier evidence of a positive relationship between inequality and performance. Moreover, as with OLS estimates, quantile regression estimates reveal higher coefficients for manufacturing firms compared to services firms.

3.3 Instrumental Variables

An important potential source of bias in OLS is simultaneity or reverse causality. For instance, under rent sharing, firms may increase their wage dispersion if high-wage workers have more bargaining power than low-wage workers. In this case, exogenous shocks that increased profits would lead to increases in wage inequality - which could be mistakenly interpreted as a positive impact of wage inequality upon firm performance.

In order to address this problem, this subsection presents results obtained through 2SLS estimates, as specified in the following equations:

$$ln(P/n)_{jt} = \beta M_{zjt} + X_{jt}'\delta + F_{jt}'\varphi + \epsilon_{jt} \tag{6}$$

$$M_{zjt} = \varpi_1 M_{zj(t-1)} + \varpi_2 M_{zj(t-2)} + \beta^* ln(P/n)_{jt} + X_{jt}'\phi + F_{jt}'\theta + \upsilon_{jt} \tag{7}$$

Similarly to Heyman (2005), lagged wage dispersion measures are used here as instrumental variables (IV). In our case, the first and second lags ($t-1$ and $t-2$) of M_{zj} were used as IV. The rationale for this choice is that lagged values of inequality will be (positively) correlated with their current values; and it can be assumed that current performance is driven only by current values of inequality (thus meeting the exclusion restriction).

Table 7 shows the value-added coefficients for 2000 and 2001 (the second stage of the 2SLS regression - equation

All coefficients of the sigma variable in Table 7 have positive and statistically significant values, regardless of the sectors or years sampled. This can be regarded as strong evidence in favor of the hypothesis of a positive impact of within-firm wage dispersion upon performance. Moreover, these results again point to higher coefficients for manufacturing compared to services.

4 Longitudinal Analysis

The previous sections showed that the cross-sectional analysis supports the tournament theory for the Brazilian case. This section conducts a longitudinal analysis, by estimating both pooled regressions and fixed-effect regressions.

4.1 Pooled Regressions

Here we pool the annual observations of firms (1998-2001) as in cross-sectional time series data and provide results that serve as a benchmark for the fixed-effects results.[8] The first specification corresponds to the pooled OLS regression:

$$ln(P/n)_{jt} = \beta M_{zjt} + X'_{jt}\delta + F'_{jt}\varphi + \lambda_t + \epsilon_{jt}, \tag{8}$$

where λ_t is a set of dummies for each year.

As with cross-sectional analyses, we consider separately all firms, manufacturing firms only and services firms only. The results for the coefficients with pooled OLS regressions are shown on the first line of Table 8. All coefficients are positive and significant at the 1% level. Moreover, the coefficient for the sample of manufacturing firms (1.23) is nearly twice that of the coefficient obtained for the sample of service firms (0.64).

The second specification is that of the quantile regression with pooled data, as shown in the following equation:

$$ln(P/n)_j = Z'_j\beta_\theta + \lambda_{tj\theta} + u_{\theta_j} \tag{9}$$

As in previous analysis, the quantile regression coefficients were obtained for the 10th, 25th, 50th, 75th and 90th percentiles. The results for the coefficients obtained with the specification above are shown in Table 8, lines 2 (P10) through 6 (P90). Again, all coefficients are positive and significant at the 1% level and the coefficient values increase at higher percentiles, independently of the sector.

The third specification is that of the pooled 2SLS regression, as shown in the following equations, in which, as before, lagged wage dispersion measures are used as instrumental

[8] The standard errors of the coefficients obtained from the pooled regressions are corrected for clusters, i.e., repeated observations for the same firm.

variables (IV):

$$ln(P/n)_{jt} = \beta M_{zjt} + X'_{jt}\delta + F'_{jt}\varphi + \lambda_t + \epsilon_{jt} \tag{10}$$

$$M_{zjt} = \varpi_1 M_{zj(t-1)} + \varpi_2 M_{zj(t-2)} + \beta^* ln(P/n)_{jt} + X'_{jt}\phi + F'_{jt}\theta + \lambda_t + \upsilon_{jt} \tag{11}$$

The results for the coefficients obtained from the pooled 2SLS specification are shown in Table 8, line 7. All the coefficients are positive and significant at 1%, regardless of the sector. Again, the coefficients obtained from the 2SLS regression yielded higher values than those obtained from the OLS estimates. 2SLS values are twice as high as those of OLS (2.63 vs. 1.23 in case of manufacturing firms, 1.45 vs. 0.64 in case of service firms and 2.32 vs. 1.08 for the whole set of firms).

4.2 Fixed Effects

A source of bias in econometric estimates that has not been addressed yet in this paper concerns unobserved heterogeneity. This section deals with this problem by using fixed effects. (Hausman's specification test was used for choosing between the random-effect and fixed-effect models, rejecting the former.) The aim here is to control for the unobserved and time-invariant heterogeneity of firms, which we do by estimating the following model:

$$ln(P/n)_{jt} = \beta M_{zjt} + X'_{jt}\delta + F'_{jt}\varphi + \lambda_t + \psi_j + \epsilon_{jt}, \tag{12}$$

where ψ_j are the firm fixed effects.

The results for the coefficient obtained from the specification of this model are displayed in Table 8, line 8. The samples of all firms and of the manufacturing firms yielded positive coefficient values (0.05 and 0.09, respectively). On the other hand, firms in the service sector had a negative coefficient value (-0.007). However, no coefficient was statistically significant.

The results above may need to be viewed with caution, since there are two issues that may affect the fixed-effect estimates: First, the panel is relatively short, with a time span of only four years, which may hinder the estimation of fixed effects. Secondly, even if the unobserved and time-invariant heterogeneity of firms is controlled for, the endogeneity problem between wage inequality and performance remains unsolved.

The solution to the first problem mentioned above implies the inclusion of further time periods in the sample. The solution to the second problem implies fixed-effect estimates with instrumental variables. Given our constraints in terms of extending the length of our sample, we consider the case of instrumental variables only.

The econometric specifications of the fixed-effect model with instrumental variables are:

$$ln(P/n)_{jt} = \beta M_{zjt} + X'_{jt}\delta + F'_{jt}\varphi + \lambda_t + \psi_j + \epsilon_{jt}, \tag{13}$$

$$M_{zjt} = \varpi_1 M_{zj(t-1)} + \varpi_2 M_{zj(t-2)} + \beta^* ln(P/n)_{jt} + X'_{jt}\phi + F'_{jt}\theta + \zeta_t + \psi_j + \upsilon_{jt}. \tag{14}$$

The results for the coefficients obtained from this model are shown in Table 8, last line. Unlike fixed-effect estimates without addressing endogeneity, here all coefficients yielded positive values. However, only the coefficient for the service sector was significant.

5 Conclusions

This paper studies the relationship between within-firm wage dispersion and performance for a large set of Brazilian firms both from the manufacturing and services sectors. Our estimates are obtained from different within-firm wage dispersion measures (conditional and unconditional inequality), different sectors of economic activity (industry and services), different performance measures (profit per worker and added value per worker), different sample characteristics (cross-sectional and longitudinal designs) and different econometric estimators (OLS, quantile regression, 2SLS, fixed effects and fixed effects with instrumental variables).

While most econometric analysis yields positive and statistically significant coefficients, fixed effects did not. However, the latter results may need to be interpreted with caution, as our panel is relatively short.

An additional aim of this study was to obtain results so as to assess the conflicting theories in this literature (tournament and fairness) for the case of Brazil. No results that could confirm the fairness hypothesis were obtained in any of the specifications or subsamples used, i.e., there was no negative and statistically significant correlation between wage dispersion and firm performance.

Finally, two sets of results deserve special attention: First, the quantile regression results lend further support to the result of a positive impact of wage inequality on firm performance. Second, overall, the results suggest that incentives generated by wage inequality are more effective in the case of manufacturing firms. According to Lazear (1989, 1995) and others, the weaker effect of inequality on the performance of service firms may be driven by a larger share of non-cooperative workers ('hawks') in this sector compared to the case of manufacturing.

References

Akerloff, G. & Yellen, J. (1988), 'Fairness and unemployment', *American Economic Review (Papers and Proceedings)* **78**(2), 44–49.

Akerloff, G. & Yellen, J. (1990), 'The fair wage-effort hypothesis and unemployment', *Quarterly Journal of Economics* **105**(2), 255–283.

DeBrock, L., Hendricks, W. & Koenker, R. (2001), 'Pay and performance: The impact of salary distribution on firm level outcomes in baseball', *Department of Economics, University of Illinois, mimeo* .

Eriksson, T. (1999), 'Executive compensation and tournament theory: Empirical tests on danish data', *Journal of Labor Economics* **17**(2), 262–280.

Heyman, F. (2005), 'Pay inequality and firm performance: Evidence from matched employer-employee data', *Applied Economics* **37**(11), 1313–1327.

Hibbs, A. & Locking, H. (2000), 'Wage dispersion and productive efficiency: Evidence for Sweden', *Journal of Labor Economics* **18**(4), 755–782.

Koenker, R. & Bassett, G. (1978), 'Regression quantiles', *Econometrica* **46**, 33–50.

Lallemand, T., Plasman, R. & Rycx, F. (2004), 'Intra-firm wage dispersion and firm performance: Evidence from linked employer-employee data', *Kyklos* **57**(4), 541–566.

Lallemand, T., Plasman, R. & Rycx, F. (2007), Wage structure and firm productivity in Belgium, NBER Working Papers 12978.

Lazear, E. (1989), 'Pay equality and industrial politics', *Journal of Political Economy* **97**, 561–580.

Lazear, E. (1995), *Personnel Economics*, M.I.T. Press, Massachusetts.

Lazear, E. & Rosen, S. (1981), 'Rank-order tournaments as optimum labor contracts', *Journal of Political Economy* **89**(5), 841–864.

Leonard, J. (1990), 'Executive pay and firm performance', *Industrial and Labor Relations Review* **43**, 13–29.

Main, B., O'Reilly, C. & Wade, J. (1993), 'Top executive pay: Tournament or teamwork?', *Journal of Labor Economics* **11**(4), 606–628.

Martins, P. S. (forthcoming), 'Dispersion in wage premiums and firm performance', *Economics Letters* .

Martins, P. S. & Esteves, L. A. (2006), Is there rent sharing in developing countries? matched-panel evidence from brazil, IZA Discussion Papers 2317.

Martins, P. S. & Pereira, P. T. (2004), 'Does education reduce wage inequality? Quantile regression evidence from 16 countries', *Labour Economics* **11**(3), 355–371.

Milgrom, P. (1988), 'Employment contracts, influence activities and efficient organisation design', *Journal of Political Economy* **96**, 42–60.

Milgrom, P. & Roberts, J. (1990), 'The efficiency of equity in organisational decision process', *American Economic Review (Papers and Proceedings)* **80**, 154–159.

Winter-Ebmer, R. & Zweimuller, J. (1999), 'Intra-firm wage dispersion and firm performance', *Kyklos* **52**(4), 555–572.

Table 1: Descriptive Statistics, All Firms

	Mean (Standard Deviation)			
	1998	1999	2000	2001
Variables				
Ln Employment	5,07	5.10	5.16	5.18
	(0.91)	(0.91)	(0.91)	(0.93)
Profits pw	934	1,022	1,051	926
	(3,000)	(3,079)	(3,017)	(2,984)
Value Added pw	16,007	15,909	16,198	16,019
	(14,305)	(14,722)	(14,813)	(14,923)
Sigma	0.16	0.15	0.15	0.15
	(0.09)	(0.09)	(0.09)	(0.10)
Standard Deviation	2.17	2.08	2.09	2.22
	(2.20)	(3.36)	(2.89)	(2.65)
Coef. of Variation	75.91	75.30	76.05	79.69
	(39.74)	(41.09)	(40.39)	(47.07)
Maxmin	37.01	38.04	38.76	45.48
	(92.33)	(82.50)	(57.45)	(70.64)
Tenure	36.41	39.44	40.76	42.84
	(21.84)	(22.61)	(22.61)	(23.52)
Schooling	6.67	6.87	7.07	7.25
	(1.77)	(1.74)	(1.76)	(1.77)
Age	32.43	32.76	32.88	33.17
	(3.75)	(3.75)	(3.79)	(3.81)
Fem	0.27	0.27	0.27	0.27
	(0.24)	(0.23)	(0.23)	(0.24)
Obs.	7689	7689	7689	7689

Notes: (1) Monetary values in R$ at 1998 prices (INPC deflator).

A Main tables

Table 2: Cross Section Analysis, OLS - All Firms

	Dependent Variable: Ln Profits pw			
	1998	1999	2000	2001
Regressors				
Sigma	0,36	0,25	0,52	0,77
	(0,26)	(0,31)	(0,25)**	(0,22)***
Standard Deviation	0,04	0,04	0,05	0,07
	(0,01)***	(0,01)***	(0,01)***	(0,01)***
Coef. Variation	0,001	0,0005	0,001	0,002
	(0,0006)**	(0,0006)	(0,0005)*	(0,0005)***
Maxmin	0,0001	0,0004	0,001	0,001
	(0,0002)	(0,0005)	(0,0004)**	(0,0004)**
Observaes	5218	5287	5297	5200

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 3: Cross Section Analysis, OLS - All Firms

	Dependent Variable: Ln Value Added pw			
	1998	1999	2000	2001
Regressors				
Sigma	1,17	1,11	1,30	1,00
	(0,07)***	(0,07)***	(0,07)***	(0,06)***
Standard Deviation	0,13	0,13	0,14	0,08
	(0,003)***	(0,001)***	(0,002)***	(0,002)***
Coef. Variation	0,002	0,001	0,001	0,001
	(0,0001)***	(0,0001)***	(0,0001)***	(0,0001)***
Maxmin	0,001	0,0007	0,0009	0,0006
	(0,00007)***	(0,00007)***	(0,0001)***	(0,0001)***
Obs.	7635	7630	7621	7626

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 4: Cross Section Analysis, OLS - Manufacturing Firms

	Dependent Variable: Ln Value Added pw			
	1998	1999	2000	2001
Regressors				
Sigma	1,33	1,23	1,45	1,19
	(0,10)***	(0,10)***	(0,09)***	(0,09)***
Standard Deviation	0,13	0,14	0,15	0,09
	(0,004)***	(0,002)***	(0,002)***	(0,003)***
Coef. Variation	0,002	0,001	0,002	0,001
	(0,0002)***	(0,0002)***	(0,0002)***	(0,0002)***
Maxmin	0,0002	0,0008	0,001	0,0007
	(0,0000)***	(0,0001)***	(0,0002)***	(0,0001)***
Obs.	4940	4944	4935	4944

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 5: Cross Section Analysis, OLS - Services Firms

	Dependent Variable: Ln Value Added pw			
	1998	1999	2000	2001
Regressors				
Sigma	0,74	0,66	0,93	0,54
	(0,08)***	(0,07)***	(0,08)***	(0,06)***
Standard Deviation	0,10	0,09	0,11	0,05
	(0,004)***	(0,003)***	(0,004)***	(0,003)***
Coef. Variation	0,001	0,0006	0,0008	0,0003
	(0,0002)***	(0,0001)***	(0,0001)***	(0,0001)***
Maxmin	0,0007	0,0003	0,0004	0,0003
	(0,0001)***	(0,0001)***	(0,0001)***	(0,0001)***
Obs.	2695	2686	2683	2681

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 6: Cross Section Analysis, Quantile Regressions - All Firms

	Dependent Variable: Ln Value Added pw					
	P10	P25	P50	P75	P90	N
Sigma						
1998	0,85	1,03	1,16	1,15	1,19	7633
	(2,53)	(1,37)	(0,96)	(0,33)***	(1,53)	
1999	0,67	0,87	1,03	1,20	1,28	7628
	(1,19)	(0,41)**	(0,48)**	(0,26)**	(0,48)***	
2000	0,75	0,90	1,29	1,45	1,49	7614
	(9,01)	(2,62)	(1,50)	(0,97)	(3,16)	
2001	0,54	0,78	0,96	1,02	1,09	7621
	(0,69)	(0,23)***	(0,28)***	(0,21)***	(0,21)***	

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 7: Cross Section Analysis, 2SLS

	Dependent Variable: Ln Value Added pw_t					
	Manufacturing 2000	Manufacturing 2001	Services 2000	Services 2001	All Firms 2000	All Firms 2001
Second Stage						
Sigma_t	2,69	2,43	1,58	1,34	2,46	2,12
	(0,39)***	(0,35)***	(0,16)***	(0,17)***	(0,25)***	(0,23)***
First Stage						
Sigma_{t-1}	0,44	0,32	0,39	0,38	0,42	0,35
	(0,01)***	(0,01)***	(0,02)***	(0,02)**	(0,01)***	(0,01)***
Sigma_{t-2}	0,21	0,29	0,24	0,25	0,23	0,28
	(0,01)***	(0,01)***	(0,01)***	(0,02)***	(0,01)***	(0,01)***
$\text{Sargan}^{(a)}$	2,00	0,45	0,19	5,13	1,64	2,01
	(0,15)	(0,50)	(0,66)	(0,02)	(0,20)	(0,15)
Shea R^2 *Partial*	0,32	0,26	0,32	0,23	0,33	0,26

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator); (a) P-Valor in parenthesis.

Table 8: Longitudinal Analysis

	Dependent Variable: Ln VA pw_t		
	Manufacturing	Services	All Firms
Pooled OLS	1,23	0,64	1,08
	(0,12)***	(0,07)***	(0,08)***
P10	0,86	0,31	0,69
	(0,16)***	(0,11)***	(0,08)***
P25	1,07	0,48	0,91
	(0,07)***	(0,05)***	(0,05)***
P50	1,22	0,63	1,08
	(0,05)***	(0,04)***	(0,03)***
P75	1,30	0,80	1,14
	(0,04)***	(0,04)***	(0,03)***
P90	1,41	0,89	1,25
	(0,05)***	(0,05)***	(0,04)***
Pooled 2SLS	2,63	1,45	2,32
	(0,26)***	(0,12)***	(0,17)***
Fixed Effects	0,09	-0,007	0,05
	(0,12)	(0,04)	(0,08)
IV - Fixed Effects	0,35	0,24	0,31
	(0,43)	(0,16)***	(0,30)

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Cluster standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location, sector, and year; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 9: List of Variables

Variables	Description:
Ln Employment	Log of employees - proxy for firm size
Profits pw	Profits per worker by firm in R$, 1998 prices
Value Added pw	Value Added per worker by firm in R$, 1998 prices
Ln Profit	Log of Profits pw
Ln Value Added	Log of Value Added pw
Sigma	Standard Error of Wage Regression by Firm-Year
Standard Deviation	Standard Deviation of Wages by Firm-Year
Coef. Variation	Wages Coefficient of Variation by Firm-Year
Maxmin	Maximum-Minimun Wages Ratio by Firm-Year
Tenure	Tenure in months - Firm-Year Average
Schooling	Years of Schooling - Firm-Year Average
Age	Age of Workers - Firm-Year Average
Fem	Female participation in % by Firm-Year
Location	Dummies for 27 Brazilian States
Sector	Dummies for sectors 3-digit CNAE

B Additional tables

Table 10: Descriptive Statistics, Manufacturing Firms

	Mean (Standard Deviation)			
	1998	1999	2000	2001
Varibles				
Ln Employment	4.94	4.97	5.02	5.03
	(0.82)	(0.81)	(0.82)	(0.83)
Profits pw	836	1,049	1,103	952
	(3,392)	(3,511)	(3,437)	(3,420)
Value Added pw	19,060	19,102	19,583	19,453
	(16,386)	(16,908)	(16,911)	(17,103)
Sigma	0.16	0.16	0.16	0.16
	(0.09)	(0.09)	(0.09)	(0.10)
Standard Deviation	2.34	2.27	2.28	2.40
	(2.34)	(3.93)	(3.36)	(2.95)
Coef. of Variation	80.05	79.45	80.17	83.24
	(40.44)	(40.33)	(38.97)	(45.87)
Maxmin	37.19	36.85	37.46	43.69
	(108.03)	(90.64)	(52.96)	(70.24)
Tenure	40.23	43.03	44.00	46.69
	(21.60)	(22.26)	(22.30)	(22.82)
Schooling	6.63	6.85	7.07	7.24
	(1.67)	(1.64)	(1.67)	(1.70)
Age	31.80	32.10	32.18	32.50
	(3.46)	(3.44)	(3.43)	(3.46)
Fem	0.28	0.28	0.28	0.28
	(0.24)	(0.24)	(0.23)	(0.24)
Obs.	4990	4990	4990	4990

Notes: (1) Monetary values in R$ at 1998 prices (INPC deflator).

Table 11: Descriptive Statistics, Services Firms

	Mean (Standard Deviation)			
	1998	1999	2000	2001
Variables				
Ln Employment	5.31	5.35	5.42	5.46
	(1.03)	(1.02)	(1.02)	(1.04)
Profits pw	1.114.52	971.80	954.80	879.64
	(2,080.11)	(2,053.32)	(2,021.02)	(1,931.38)
Value Added pw	10,563	10,005	9,941	9,669
	(6,302)	(5,935)	(6,009)	(5,614)
Sigma	0.15	0.14	0.14	0.15
	(0.09)	(0.09)	(0.09)	(0.09)
Standard Deviation	1.84	1.73	1.75	1.90
	(1.87)	(1.87)	(1.67)	(1.95)
Coef. Variation	68.30	67.61	68.45	73.12
	(37.23)	(41.38)	(41.85)	(48.56)
Maxmin	36.68	40.24	41.18	48.79
	(52.40)	(64.78)	(64.88)	(71.28)
Tenure	29.40	32.81	34.78	36.62
	(20.50)	(21.75)	(22.61)	(23.53)
Schooling	6.73	6.92	7.09	7.27
	(1.94)	(1.92)	(1.92)	(1.90)
Age	33.60	34.00	34.17	34.40
	(3.97)	(4.00)	(4.07)	(4.11)
Fem	0.25	0.25	0.25	0.26
	(0.23)	(0.23)	(0.23)	(0.25)
Obs.	2699	2699	2699	2699

Notes: (1) Monetary values in R$ at 1998 prices (INPC deflator).

Table 12: Cross Section Analysis, Quantile Regressions - Manufacturing Firms

	Dependent Variable: Ln Value Added pw					
	P10	P25	P50	P75	P90	N
Sigma						
1998	0,80	1,24	1,43	1,36	1,45	4938
	(0,71)	(0,40)***	(0,15)***	(0,06)***	(0,21)***	
1999	0,64	0,86	1,07	1,34	1,57	4942
	(0,87)	(0,19)***	(0,22)***	(0,14)***	(0,23)***	
2000	1,15	1,16	1,39	1,46	1,65	4934
	(0,65)*	(0,38)***	(0,15)***	(0,15)***	(0,38)***	
2001	1,04	1,13	1,14	1,26	1,34	4940
	(0,60)*	(0,17)***	(0,18)***	(0,21)***	(0,11)***	

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

Table 13: Cross Section Analysis, Quantile Regressions - Services Firms

	Dependent Variable: Ln Value Added pw					
	P10	P25	P50	P75	P90	N
Sigma						
1998	0,63	0,63	0,68	0,79	0,76	2695
	(0,30)**	(0,12)***	(0,08)***	(0,10)***	(0,14)***	
1999	0,32	0,58	0,60	0,75	0,83	2686
	(0,17)*	(0,08)***	(0,06)***	(0,09)***	(0,11)***	
2000	0,25	0,59	0,91	1,01	1,20	2680
	(0,23)	(0,11)***	(0,09)***	(0,07)***	(0,14)***	
2001	0,16	0,23	0,55	0,67	0,84	2681
	(0,12)	(0,09)**	(0,08)***	(0,06)***	(0,07)***	

Notes: (1) Significant at 1% (***), 5 % (**), and 10% (*); (2) Robust standard errors in parenthesis; (3) Each cell corresponds a different regression; (4) Regression also include the following control variables: ln employment, schooling, tenure, % female, dummies for location and sector; (5) Monetary values in R$ at 1998 prices (INPC deflator).

C Theory

Akerloff & Yellen (1988) presents an efficiency wage model where, due to imperfect monitoring and imperfect measurement of worker heterogeneity, the problem of the firm lies in managing incentives in order to elicit more effort from workers. In this study, the authors suggest that the workers' effort function can be expressed as:

$$e = e(\sigma^2(w)), \tag{15}$$

where e is the effort, w is the wage and σ^2 is the within-firm wage differential. The expression above suggests that workers' effort does not depend solely on wages, but also on within-firm wage dispersion. In this case, higher wages and lower within-firm wage dispersion allow employers to demand more effort from their workers.

The idea of a tradeoff between wage dispersion and firm performance originating from a feeling of fairness is elaborated in Akerloff & Yellen (1990), where they assume that workers compare their wages internally and externally and choose their effort based on:

$$e = min\left[\frac{w}{w_f}, 1\right], \tag{16}$$

where w is the wage earned by the worker, w_f is the wage level considered to be fair and 1 is the normal effort level. The equation above shows that workers' effort falls short of the wage they regard as fair. The idea of a fair wage, as previously mentioned, is based on the comparison of wage differentials within and outside firms.

A second theoretical argument that supports a negative relationship between within-firm relative wage dispersion and economic performance is developed in Milgrom (1988) and Milgrom & Roberts (1990). The authors argue that white-collar workers have incentives to withhold information from managers in order to increase their influence and to engage in rent-seeking activities instead of productive work. The authors also state that the implementation of some degree of wage equity can reduce the tendency of white-collar workers to take personal interest decisions instead of decisions that may be profitable for the organization as a whole.

On the other hand, Lazear & Rosen (1981) points to the benefits of a more dispersed wage structure derived from a performance-based pay system, in which the most productive workers are awarded the largest prize; therefore, a higher effort is demanded from workers so that they remain in contention for the prize (bonuses or promotions). This idea can be expressed as follows: consider two identical risk-neutral workers, j and k, and a risk-neutral firm with a compensation scheme in the value of W_h for the most productive workers and of W_l for the least productive ones, where $W_h > W_l$. The output level of each player is given according to the following equation:

$$q_i = e_i + \epsilon_i, \tag{17}$$

where q is the output level, e is the effort level, $i = [j, k]$ and ϵ is a random component of output (e.g.: luck). Let us now suppose that the utility expected from player j is given by the following equation:

$$U_j = P\left(W_h - C(e_j)\right) + (1 - P)\left(W_l - C(e_j)\right), \tag{18}$$

where U is the expected utility, P is the probability to win the game, and $C(.)$ is a cost function with $C' > 0$ and $C'' > 0$. The probability of player j winning the game is:

$$prob\left(q_j > q_i\right) = prob((\epsilon_k - \epsilon_j) < (e_j - e_k)) = prob((e_j - e_k) > \zeta) = G\left(e_j - e_k\right), \tag{19}$$

where $\zeta = (\epsilon_k - \epsilon_j)$, ζ distributed in $g\left(\zeta\right)$ with mean zero, and G is the cumulative density function of ζ. The worker maximizes his expected utility, U, by choosing an effort level in which the marginal cost of his effort equates to the marginal benefit, that is:

$$\left(W_h - W_l\right)\frac{\partial\, P}{\partial\, e_i} - \frac{\partial\, C}{\partial\, e_i} = 0. \tag{20}$$

If both players, j and k, maximize their utilities, the marginal probability to win relative to their effort will be given by:

$$\frac{\partial\, P}{\partial\, e_j} = \frac{\partial\, G\left(e_j - e_k\right)}{\partial\, e_j} = g\left(e_j - e_k\right). \tag{21}$$

Substituting the equation above into the first-order condition for the maximization of the expected utility yields:

$$\left(W_h - W_l\right) g\left(e_j - e_k\right) = \frac{\partial\, C}{\partial\, e_j}. \tag{22}$$

Given the hypothesis of homogeneous work, both players will choose the same effort level. According to the symmetric Nash equilibrium $e_j = e_k$, and the outcome of the game is random with $P = 0.5$. Thus, the latter equation can be written as:

$$\left(W_h - W_l\right) g\left(0\right) = \frac{\partial\, C}{\partial\, e_j}. \tag{23}$$

The equation above indicates that a higher wage dispersion deriving from output-based reward strategies, $(W_h - W_l)$, implies larger marginal costs of effort for players, or simply higher levels of effort, since $C' > 0$ e $C'' > 0$, as previously stated.

Lazear (1989, 1995) also argue that the efficiency of an output-based compensation system can be offset (or even neutralized) by the effect of a lower level of work cohesion due to non-cooperative behaviors. They show that high wage compression is crucial for an organization in which most workers are non-cooperative (hawks). According to these authors, following the arguments of Milgrom (1988) and Milgrom & Roberts (1990), hawks are commonly found among management and supervision staff members and among white-collar workers.

D RAIS

RAIS ('Relação Anual de Informações Sociais', Annual Social Information Report) is an administrative report filed by all tax registered Brazilian establishments. Since the information may be used for investigation about labor legislation compliance, firms that do not comply with it do not file in RAIS. Thus, this data set can be considered a census of the formal Brazilian labor market (State-owned enterprises, public administration and non-profit organizations are also required to file the report.) Firms that do not provide accurate information will be committing an offense sanctioned by law, a threat that is likely to lead to very high standards of data quality.

RAIS covers the whole country and is carried out annually. The information is collected every year in the first quarter, referring to the previous year. Every tax registered enterprise receives a unique tax number (CNPJ). This number is composed by a specific firm part and a complement for each unit (local plant or establishment) that the firm operates.

The main variables available from the survey at the establishment level are:

- Geographic location: State, metropolitan region, county;
- Activity sector: CNAE (National Economic Activity Classification); sector Level (10 categories); activity (42 categories); sub-activity (about 560 categories);
- Establishment Size: number of workers, number of wage earners, number of owners;
- Establishment Type: Private enterprise, private foundation, State-owned enterprise, State foundation, joint public-private enterprise, non-governmental organization, government, nonprofit enterprise, notary.

At the employee level, the following information is available (although we did not obtain access to all variables listed):

- Occupation: occupation classes (CBO-Brazilian Occupation Classification system - about 350 categories); subgroup (84 categories); group (11 categories);
- Personal Characteristics: schooling (9 classes), age, gender, nationality.
- Contract Information: month of admission, month of separation, December wage rate (13th monthly salary), average yearly wage, tenure, separation cause (fired with/without fair reason, separation with/without fair reason, retiring, transfer to other units or firm), contract type (work card, civil service, isolated worker, temporary worker), contract status (in activity or paid leave, leave without paid, occupation accident, military service, maternity leave, sick leave, inactive), admission type (first placement in firm, re-employment, transferred), contract hours (exclusive overtime).

As some other matched employer-employee panels, RAIS is based on worker spells, defined by an occupation-establishment-contract group in each year. In other words, if a worker changes his/her occupation or establishment or contract type in a given year, there will be one separate observation for each case.

With the establishment identification number (CNPJ) it is possible to follow all establishments that file the RAIS survey. Moreover, with the worker's national insurance number, it is possible to follow all workers that remain in the formal sector and to match the worker's characteristics with those of the establishment. Therefore, we can create a panel that matches workers to their establishments and follow each of them over time. It was using the firm identification numbers that we have merged the three data sets described in this appendix.