EG Estudios Gerenciales
Journal of Management and Economics for Iberoamerica

Research article

The more intra-firm equity, the higher employee retention: Evidence from the universe of urban firms in Colombia

Eduardo Lora
Associate researcher, Fedesarrollo, Bogotá, Colombia.
Eduardo.A.Lora@gmail.com

Abstract
Fairness and equity considerations may influence workers' decision to stay or quit their jobs. This paper tests several hypotheses derived from theory on the relation between intra-firm equity and employee retention. It uses administrative data covering the universe of urban firms in Colombia between 2008 and 2016. The average wage of the firm, the intra-firm dispersion of real wage adjustments, and the share of women in the firm are significantly associated with retention rates in the directions predicted by theory. The intra-firm wage gap is directly associated with retention, at odds with theory. The results suggest that equity considerations are important in employees' decision to stay or quit.

Keywords: intra-firm equity; employee retention; personnel turnover; wage structure.

A mayor equidad en las empresas, mayor retención del personal: Evidencia del universo de empresas urbanas en Colombia

Resumen
Las consideraciones de equidad pueden influir en la decisión de los trabajadores de permanecer o renunciar a sus trabajos. Este artículo plantea varias hipótesis derivadas de la teoría sobre la relación entre la equidad intraempresa y la retención de personal, y utiliza datos administrativos del universo de empresas urbanas en Colombia entre 2008 y 2016. El salario promedio de la empresa, la dispersión intraempresa de los ajustes salariales y la participación de las mujeres en la empresa están significativamente asociados con las tasas de retención, como lo predice la teoría. La brecha salarial intraempresarial está directamente asociada con la retención, en contra de la teoría. Así, las consideraciones de equidad parecen influir en la decisión de los empleados de quedarse o renunciar.

Palabras clave: equidad intraempresarial; retención de personal; rotación de personal; estructura salarial.

Mais equidade nas empresas, maior retenção de pessoal: evidências do universo das empresas urbanas na Colômbia

Resumo
As considerações de equidade podem influenciar as decisões dos trabalhadores de permanecerem ou abandonarem os seus empregos. Este artigo propõe diversas hipóteses derivadas da teoria sobre a relação entre patrimônio intra-empresa e retenção de funcionários, e utiliza dados administrativos do universo de empresas urbanas na Colômbia entre 2008 e 2016. O salário médio da empresa, a dispersão intra-empresa dos reajustes salariais e da participação das mulheres na empresa estão significativamente associados às taxas de retenção, conforme prevê a teoria. A disparidade salarial intra-empresa está diretamente associada à retenção, contrariamente à teoria. Assim, as considerações de equidade parecem influenciar a decisão dos empregados de permanecer ou demitir-se.

Palavras-chave: patrimônio intra-empresa; retenção de pessoal; rotação de pessoal; estrutura salarial.

* Corresponding author.

JEL classification: J31; J33; J63.

How to cite: Lora, E. (2023). The more intra-firm equity, the higher employee retention: Evidence from the universe of urban firms in Colombia. *Estudios Gerenciales, 39*(168), 327-340. https://doi.org/10.18046/j.estger.2023.168.5897

DOI: https://doi.org/10.18046/j.estger.2023.168.5897

Received: 14-12-2022
Accepted: 20-04-2023
Available on line: 30-09-2023

UNIVERSIDAD ICESI

1. Introduction

Employee turnover is costly to firms because personnel selection, recruitment, and training demand expenditures and organizational efforts. Especially for firms in sophisticated industries, employee turnover is additionally costly because the productivity of any team depends on the quality of the matching between all its members (Kremer, 1993; O'Connell, and Mei-Chuan Kung, 2007). High employee retention facilitates training, both at the individual and the team level, given that any investment is more profitable the longer the period over which returns can accrue. Therefore, firms facing high employee turnover may be discouraged to invest in their own personnel, falling into a low productivity trap (Glance, Hogg, and Huberman, 1997).

Many critical factors identified by motivation theories may influence workers' decision to stay or quit their jobs. Among them are fairness and equity considerations, since employees want to be treated and rewarded in a fair and equitable manner (Ramlall, 2004). Abundant studies for specific firms, industries and/or professions in both developed and developing economies have identified factors affecting employee retention (see the review by Das and Baruah, 2013). With very few recent exceptions (Pfeifer and Schneck, 2012; Mohrenweisen and Pfeifer, 2019), employee retention studies have not taken advantage of the growing availability of matched employer-employee panel data that provide information on all registered firms and their workers in a location over several years. Although this type of data does not contain the nuanced information by individual often used in psychology, business, and managerial studies, it does provide enough information at the individual level to compute precise employee retention measures, along with indicators on wage levels, wage gaps, wage adjustments, and gender equity across large numbers of firms and industries over time.

The purpose of this paper is to assess the influence of equity variables on employee retention; these variables may be quantified with a rich matched employer-employee database for Colombia over 2008-2016. Apart from the availability of data, Colombia is a fitting case to study the influence of intra-firm equity variables on personnel retention for several reasons. First, Colombia's labour earnings are the third most concentrated among 38 countries in the world (Rodríguez-Castelán et al., 2016). Second, Colombia is the Latin American country with the largest share of salaried workers with less than five years of tenure (71.8 percent versus 51.8 percent on average in the whole region between 2010 and 2019, according to the Inter-American Development Bank, 2020). And third, most firms are poorly managed: Colombia is ranked 26th among the 34 countries included in the World Management Survey, which assesses the quality of management practices in large firms based on criteria derived from profit maximization theory and empirical evidence (Scur et al., 2021). Colombian firms have the greatest deficiencies in human resources management. According to the Survey, good human resources management implies promoting and remunerating employees according to their performance and contribution to the success of the company, being very selective in the hiring of new staff and very determined to retain the best employees. Very few Colombian companies rate well in these matters (Bloom et at., 2012).

The dependent variable in this study is the employee retention rate by firm, city, and year. Based on equity theory, relative deprivation theory, efficiency wage theory, and gender theory, the main explanatory variables are measures of equity in remuneration, such as firms' average wages relative to those paid by comparable firms, intra-firm wage gaps, intra-firm real average wage adjustments and their variance; and gender equity measures, such as the share of women and the gender wage gap within firms. Since the dependent variable may be correlated with unobserved firm characteristics, standard estimators such as ordinary least squares yield inconsistent estimates of the parameters relating the dependent and the explanatory variables. Furthermore, if the autoregressive parameter of the dependent variable is too large, the standard dynamic panel estimator (Arellano and Bond, 1991) can perform poorly. Therefore, this study uses the dynamic system estimator developed by Blundell and Bond (1998), which is suitable to deal with potential endogeneity of the explanatory variables in panel data with few periods and many observations.

The econometric results are consistent with some predictions of equity theory, efficiency wage theory, relative deprivation theory, and gender theory on the testable factors affecting employee retention. Employee retention is higher in firms that pay higher wages with respect to those paid by other firms in the same industry and city, that adjust wages uniformly, and in those where a higher proportion of their employees are women. These results are consistent with theory. However, employee retention is higher in firms with larger wage gaps between their top and their middle-wage employees, which is at odds with relative deprivation theory. Incidentally, this paper uncovers evidence that supports a prediction of labour search theory, i.e., industry concentration at the local level is positively correlated with personnel retention. Most of the results are consistent across groups of firms by size, by the size of the city where they operate, and by sector.

The rest of this paper is organized as follows. Section 2 is a short survey of the empirical literature on the determinants of employee turnover/retention, aimed at providing background on the attention given by previous studies to fairness considerations, social preferences for equity and gender differences. Section 3 elaborates on the choice of explanatory variables and the resulting testable hypotheses based on equity theory, efficiency wage theory, relative deprivation theory and

gender theory. The database and the main variables are described in Section 4. The main econometric results are presented in Section 5. Section 6 discusses the results and concludes the paper.

2. Literature background

Since the 1980s, in the US and other developed countries, personnel retention has been an important concern in sectors such as the armed forces and the health sector. Several organisational studies have addressed the issue with the aim of recommending improvements in recruitment, training, managing, and remunerating practices to increase retention in those sectors (Berger and Boyle, 1992; Bergman, Eckerling and Golander, 1984; Doig and Beck, 2005; Faris, 1984; Gilroy, Horne, and Smith,1991; Orrick, 2008).

There is a wide literature on management studies that has analyzed the determinants of turnover intentions. Individual and organizational factors have been found to affect variables like commitment, job satisfaction, personal agency, self-esteem, and others, which in turn influence turnover intentions (Jha, 2009; Flint, Haley, and McNally, 2013). The same mechanisms and factors that affect employee's turnover intentions have been considered to study actual turnover in longitudinal studies. An early review of the research on the topic concluded that intentions were more predictive of attrition than overall job satisfaction, satisfaction with work itself, or organizational commitment (Steel and Ovalle, 1984). A study of staff nurses in Taiwan (Chen, Chu, Wang, and Lin, 2008) found that, except for workload, those mechanisms and factors do not help to predict actual turnover. This suggests that a more comprehensive selection of turnover factors may be needed to explain variations in actual turnover. For instance, as predicted by social exchange theory (which states that voluntary actions initiated by employers in support of workers elicit reciprocation; see Blau,1964), and contrary to human capital theory predictions, participation in general training programs sponsored by the firm reduces turnover significantly (Koster, De Grip and Fouarge, 2011).

A few recent studies have focused on the role of intra-firm equity variables in personnel turnover. Using matched employer-employee data for Germany, Pfeifer and Schneck (2012) find that the probability of workers to quit is negatively associated with their wage as well as with an internal reference wage measure that reflects their relative wage. In a field experiment with an information treatment among employees of the University of California, Card et al. (2012) find that workers who are made aware of having relatively low wages feel less satisfied with their jobs and are more likely to search for a new job. Using an employer-employee longitudinal dataset for Germany, Mohrenweisen and Pfeifer (2019) find that intrafirm wage structure affects perceived wage fairness and turnover intentions. Dube et al. (2019), using personnel data of a large US retailer and a regression-discontinuity design, find strong causal effects of workers' own wages as well as their peers' wages on quits, suggesting that peer comparisons and fairness concerns affect personnel turnover. Their estimates suggest that when a wage raise is uniform across peers within the firm, the rate of job separations remain unchanged, and that the overall effect of wages on separations is driven mostly by peer comparisons. The study by Breza, Kaur, and Shamdasani (2018) is closely related to Dube et al (2019); they randomized Indian manufacturing workers to pay units where pay raises resulted in either compressed or unequal wage structures; in the latter case, absenteeism increased very strongly. This growing literature consistently points towards the importance of intra-firm equity variables in personnel retention.

The topics of personnel retention and turnover are related to personnel economics, which uses economic theory and empirical methods to understand the internal workings of the firm (Grund, Bryson, Dur, Harbring and Lazear, 2017). The main original tenet of personnel economics theory was that firms gravitate towards equilibrium because of the process of utility and profit maximization by workers and firms. However, personnel economists now consider firm managers and workers not to be purely egoistic "homo oeconomicus" but "homo reciprocans" (Backes-Gellner et al., 2008; Dohmen et al., 2009; and Dohmen, 2014). The new approach extends personnel economics to include fairness considerations, social preferences for equity or equality, and gender differences in decisions and behaviour.

3. Choice of explanatory variables

In this study, the choice of equity variables as explanatory variables for personnel retention rests on equity theory, efficiency wage theory, relative deprivation theory, and gender theory. Equity theory (Adams, 1965; Homans, 1974; Walster, Walster, and Berscheid, 1978) assumes that individuals judge the fairness of their relationships within organizations by comparing the balance between the inputs they contribute (e.g., work effort) and the outcomes they receive vis-à-vis the balance of inputs and outcomes of their reference groups, both inside and outside their organizations. Workers feel that equity exists when their own ratio of inputs to outcomes is like that of their comparative referents. If they perceive inequity, they attempt to reduce the distress by (1) changing their perceptions of either their own or their reference group's inputs and outcomes; (2) altering their own balance of inputs (e.g., decrease their work effort) and outcomes (e.g., get a pay raise); or (3) leaving their organizations. Based on equity theory, this study posits that personnel retention may be affected by several intra-firm variables—average wage raise (positively), wage gaps, and wage raise dispersion (both negatively)— and by one inter-firm variable, namely the average wage of the firm relative to the average wage of other firms in the same location and industry (positively).

The efficiency wage hypothesis in labor economics poses that managers have the incentive to pay their employees more than the market-clearing wage to reduce the costs associated with employee turnover, especially in sectors in which the costs of replacing labor are high (Stiglitz, 1974; Schlicht, 1978). This strategy, which is compatible with profit maximization, implies that employee retention is higher in firms that pay higher wages than other firms operating in the same labor market. Therefore, the efficiency wage hypothesis points in the same direction of equity theory.

Relative deprivation theory states that individuals in the lower strata of an organization compare the rewards that they (or their groups) receive to those received by upper-strata groups. These interclass comparisons result in feelings of injustice when individuals find that they have received less than they deserve (Martin, 1981; Crosby, 1984; Cowherd and Levine, 1992). Based on relative deprivation theory, this study posits that a wider wage gap between top and middle level employees in a firm should reduce personnel retention.

Several theories have been proposed to understand the role of gender in organizations (Ely and Meyerson, 2000). The choice of gender related variables in the present study, which is constrained by the database, cannot make justice to any of those theories. Just two intra-firm variables are considered: the share of female employees and the gender wage gap (male versus female). The proportion of females in groups has been found to be strongly correlated with a general "collective intelligence factor" that measures the performance of groups in a variety of tasks that is left unexplained by the individual intelligence of its members (Woolley et al., 2010). The effect is explained in part by the higher social sensibility of women, which improves processes. Research on the influence that gender composition variables may have on personnel retention is very limited; some (weak) evidence suggests that greater gender diversity reduces turnover intentions in the public sector (Nielsen and Madsen, 2017). The hypothesis in the present study is that a larger feminine participation is positively associated with personnel retention. As mentioned, based on equity theory, it is also hypothesized that a wider gender wage gap is negatively related with personnel retention (but causality may run in both directions because gender differences in worker turnover may incentivize employers to pay men and women differently; see Barth and Dale-Olsen, 2009). It is also hypothesized that the influence of the share of women in the firm and the gender wage gap on personnel retention may reinforce each other. A negative sign should be expected for the interaction of the two variables because a larger gender wage gap—which is expected to have a negative effect on retention—will affect workers' motivation more in firms with a larger share of women.

Among the additional controls that will be included in the estimations, there is a measure of industry concentration, which warrants some theoretical justification. Monopsony in the labour market is a recently revitalized concept in economics due to labour search theory. Within a search model, each single firm or establishment operating in a location faces its own labour supply curve because workers quit endogenously and must be replaced by new hires (Burdett and Mortensen, 1998; Manning, 2003). Workers quit more often and can be replaced by firms more easily when there are many firms of the same, or similar, industry in the location. Therefore, it is expected that industry concentration at the city level is negatively correlated with personnel turnover and positively correlated with personnel retention.

4. Data and descriptive statistics

The database is the social security administrative data collected by the Health and Social Protection Ministry, known as PILA *(Planilla Integrada de Liquidación de Aportes)*. It collects information reported monthly by firms of all sectors and sizes that contribute to the social security system. It provides information of age, sex, days of work, monthly-equivalent wage, and municipality of the establishment (firms may operate in several municipalities) of each worker in a firm. This study uses PILA data from 2008 to 2016, which were cleaned and processed by Harvard's Center for International Development to construct the Atlas of Economic Complexity of Colombia (DatlasColombia.com)[1].

The unit of observation in this paper is firm *j* that operates in city *c* in year *t*. Most firms operate in just one city. When a firm has several establishments in a city, they are consolidated into a single observation. A city may be a single municipality or a metropolitan area with population of at least 50,000 people. A metropolitan area is defined as a group of municipalities with regular commuting flows of workers among them, according to the algorithm developed by Duranton (2015). The 62 resulting municipalities are classified in four groups by size, as explained below. The industry of operation in the firm corresponds to the 4-digit ISIC code (there are 433 industries in the database). The industries are classified in eight sectors, as shown below.

The universe of firms in PILA covers firms of all sizes, but this study includes only those with an average "effective size" over the years (between 2008 and 2016) of at least 10 employees. Also, only firms that reported to PILA in 2016 and at least one previous consecutive year were included (this was done to exclude firms that closed, therefore terminating jobs unilaterally)[2]. In the computation of effective size, each employee *i* counts in proportion to the number of weeks they worked in firm *j* in city c during year *t*:

1 As inconsistencies were found by CID in the industry code of the ISIC reported by the firms, they developed a verification methodology based on information on the activities of each firm according to the Customs Office (DIAN), the Superintendence of Companies and the records of the Chambers of Commerce throughout the country. Therefore, the ISIC codes used here are not the ones reported by the firms, but those corrected by CID.

2 The database does not inform if quits are voluntary or forced. However, given the high firing costs imposed by the labour code, firing employees is not a common practice among Colombian firms.

$$effective\ size_{j,c,t} = \sum_{i \in j}\ number\ of\ weeks_{\ i,c,t} * 7/365 \quad (1)$$

Effective size as a control variable is expressed in logs. Firms are classified by size based on their effective size.

The dependent variable is the *employee retention rate* of firm *j* in period *t*, which is defined as the proportion of employees *e* that were in the firm's roster in period *t*, who appear again in the roster in period *t*+1:

$$employee\ retention\ rate_{j,c,t} = \frac{\sum_{i \in j,(t+1) \cap t} e_{i,t+1}}{\sum_{i \in j,t} e_{i,t}} \quad (2)$$

Table 1 presents descriptive statistics of the dependent variable for the 334,698 observations and information of all the explanatory variables in the regressions below. On average, employee retention rate is 0.675, with a standard deviation of 0.202. Very similar averages and standard deviations are found across the firm size, city size, and sector categories.

Table 1. Employee retention rates: descriptive statistics

	Mean	Standard deviation	Number of observations
Urban firms with at least 10 employees	0,6746	0,2022	334.698
By firm size (by effective number of employees):			
10 to 25 employees	0,6784	0,2033	175.316
25 to 50 employees	0,6665	0,1989	73.625
50 to 100 employees	0,6671	0,2010	41.747
100 or more employees	0,6798	0,2039	44.010
By city size (by working age population):			
Largest 6 cities	0,6759	0,1981	239.126
Medium-large 15 cities	0,6719	0,2106	69.217
Medium-small 16 cities	0,6556	0,2196	17.108
Smallest 25 cities	0,6943	0,2064	9.247
By industry quartiles of complexity:			
First quartile (110 industries)	0,6657	0,2038	122.845
Second quartile (110 industries)	0,6793	0,2057	61.494
Third quartile (110 industries)	0,6993	0,1864	50.648
Fourth quartile (110 industries)	0,7026	0,1842	32.023

Source: own calculations from Ministry of Health's PILA 2008-2016. It includes the observations of firms present in 2016, irrespective of their year of creation, which have data for all the explanatory variables of regression (4) in Table 4. An observation is a firm in a city in a year. See in text the definitions of variables to define the groups.

The indicators of intra-firm equity are defined as follows:

- The firm's *average wage relative to the rest of firms in the same industry and city* is the difference between the (log of) average monthly wage of the firm in city *c* in period *t* and that of the rest of the industry in the same city and period.
- *Wage gap* of firm *j* in period *t* is the (log) difference between the wage of the worker(s) at the bottom of the 90th percentile of the firm's wage distribution and the wage of the worker(s) at the bottom of the 50th percentile.
- *Average real wage adjustment* of firm *j* between period *t*-1 and period *t* is the nominal wage increase of those workers that appear in both periods in the firm workers' roster, corrected by the increase in the official consumer price index (published by DANE: the National Statistical Office).
- *Dispersion of real wage adjustments* of firm *j* in period *t* is the standard deviation of the real wage adjustments between periods *t*-1 and t of those workers that appear in both periods.
- *Share of women in personnel* is computed for all employees in the firm's roster in year *t* (irrespective of their dedication).
- *Gender wage* gap is the (log) difference between the (simple) average of the men and women wages.
- *Interaction of the last two variables (Share of women in personnel* X *Gender wage gap)*.

Variables related to the firm's industry *k* are as follows:

The normalised *Herfindahl index* of firm concentration by industry *k* in city *c* in year *t*:

$$Herfindahl\ index_{k,c,t} = \frac{\sum_{j \in k} s^2_{j,c,t} - 1/N_{k,c,t}}{1 - 1/N_{k,c,t}} \quad (4)$$

where:

$$S_{j,c,t} = \frac{effective\ size_{j,c,t}}{\sum_{j \in k} effective\ size_{j,c,t}} \quad (5)$$

and $N_{k,c,t}$ is the number of firms in industry *k* in city *c* in year *t*.

- The *Bartik instrument of nation-wide sectoral shocks* of industry *k* in city *c* in year *t* is a measure of the shocks affecting the industry nation-wide, as reflected in the change of its employment level. As suggested by Bartik (1991), to mitigate endogeneity, it is computed for each industry and city by subtracting from national employment, by industry, the employment in that industry in the city, and expressed in logs:

$$Bartik\ instrument_{k,c,t} = \log(effective\ size_{k,t} - effective\ size_{k,c,t}) - \log(effective\ size_{k,t-1} - effective\ size_{k,c,t-1}) \quad (6)$$

Finally, there is one explanatory variable related to the city *c* where the firm operates:

- *Government expenditure per capita*, in constant prices, computed from expenditure data by municipality by the National Planning Department, population by municipality and the official CPI, both by DANE.

Three categorical variables are used to explore possible sources of heterogeneity by groups, as follows:

- Firm size groups (by the average effective number of employees of the firm throughout the whole period of analysis). Firms smaller than 10 employees are not included:
  - *10* to *25* employees
  - *25* to *50* employees
  - *50* to *100* employees
  - *100* or more employees.
- City size groups by working age population (in 2016, from DANE):
  - *Largest cities:* more than 700,000
  - *Medium-large cities:* between 200,000 and 700,000
  - *Medium-small cities:* between 100,000 and 200,000
  - *Smallest cities:* less than 100,000.
- Sectors: using their 2-digit ISIC, firms are classified in sectors as follows (agriculture and mining firms are not included in this classification)
  - *Manufacturing:* ISIC codes 15-37
  - *Construction:* ISIC code 45
  - *Commerce:* ISIC codes 50-58
  - *Finance:* ISIC codes 65-67 and 70
  - *Utilities and transportation:* ISIC codes 40-43 and 60-64
  - *Services to firms:* ISIC codes 71-74
  - *Government and social services:* ISIC codes 75-85
  - *Other services:* ISIC codes 86-99.

Table 2 presents descriptive statistics of the explanatory variables that will be included in the regressions (334,698 observations). Correlations with the dependent variable are also shown. Table 3 shows the largest pairwise correlations between explanatory variables (all the correlations larger than 0.2 in absolute value are presented).

Table 2. Descriptive statistics of explanatory variables.

	Mean	Standard deviation	Minimum	Maximum	Correlation with dependent variable
Firm level variables					
Effective size (logs)	3,4084	1,1458	-3,9540	9,7669	0,0746
Average wage relative to rest of industry in city (logs)	-0,0469	0,4594	-2,3551	2,7572	0,2479
Wage gap (90-to-50 percentile ratio)	0,5967	0,4024	0,0000	4,1529	0,2048
Average real wage adjustments (logs)	0,0358	0,1046	-2,4534	2,2028	0,0601
Dispersion of real wage adjustments	0,1336	0,0894	0,0000	2,0882	-0,2247
Share of women	0,4123	0,2378	0,0020	0,9982	0,1010
Gender diversity	0,7430	0,2529	0,0070	1,0000	0,1066
Gender wage gap (logs)	0,0481	0,2385	-3,6423	4,2239	0,0489
Industry and city level variables					
Herfindahl concentration index of industry in city	0,1388	0,2130	0,0006	1,0000	0,0662
Bartik instrument of nation-wide industry shocks (logs)	0,0936	0,2748	-8,5916	14,4374	0,0177
Government expenditure per capita (constant prices, logs)	0,1606	0,3662	-1,3164	5,7390	-0,0323

Source: own calculations from Ministry of Health's PILA 2008-2016. Number of observations is 334,698 (an observation is a firm in a city in a year).

Table 3. Largest pairwise correlations between explanatory variables.

		(a)	(b)	(c)	(d)	(e)	(f)
(a)	Average wage relative to rest of industry in city (logs)		0,43				
(b)	Wage gap (90-to-50 percentile ratio)	0,43		0,22		0,30	
(c)	Dispersion of real wage adjustments		0,22				
(d)	Share of women					0,43	0,24
(e)	Gender diversity		0,30		0,43		
(f)	Gender wage gap (logs)				0,24		

Source: own calculations from Ministry of Health's PILA 2008-2016. An observation is a firm in a city in a year. Correlations computed for the 334,698 observations with data for all variables. All the correlations larger (in absolute value) than 0.2 are shown.

5. Econometric estimates

The database is unbalanced and has large cross-sectional variation (61,887 firms) with few years of time variation (2008-2016). The fact that the lagged dependent variable may be correlated with unobserved city-level characteristics makes standard estimators, such as ordinary least squares, inconsistent. The Arellano and Bond (1991) estimator is designed to deal with this problem based on the assumption of sequential exogeneity: current unobserved determinants of employee retention (the error term) are not correlated with past realizations of employee retention (lagged values of the dependent variable). It uses moment conditions in which lagged levels and lagged differences are instruments for the endogenous variables in the level equation and it can deal with unbalanced panels, such as our database. However, the Arellano and Bond estimator can perform poorly if the autoregressive parameter is too large or the ratio of the variance of the panel-level effect to the variance of idiosyncratic error is too large. To deal with this problem, the Blundell and Bond (1998) dynamic system estimator uses additional moment conditions, assuming that there is no autocorrelation in the idiosyncratic errors and that the city-level fixed effect is uncorrelated with the first difference of the first observation of the dependent variable. The core model for all Blundell and Bond regressions below is the following:

$$y_{j,c,t} = \beta_0 + \beta_1 y_{j,c,t-1} + \beta_2 X_{j,c,t} + \beta_3 T_t + \theta_{j,c} + u_{j,c,t} \quad (7)$$

where subscript *j* refers to firms, subscript *c* refers to cities, and subscript *t* refers to years. The employee retention rate *y* is the dependent variable, and it has an autocorrelation component. All other observed variables are included in vector *X*. They may be endogenous to the other variables or predetermined (that is, autocorrelated but independent from the other variables). Vector *T* is the set of year dummies, which are exogenous. Vector θ captures all firm fixed effects (which are time invariant, and which make redundant other firm-and-time invariant fixed effects, such as city or industry). Finally, *u* is the error term. All the β coefficients are estimated with the Blundell and Bond dynamic system estimator (which differentiates between exogenous, predetermined, and endogenous regressors).

All the variables in vector *X* that vary across *all* firms in a city and year are considered endogenous. The remaining variables, which do not vary across *all* firms in a city and year (either related to the firm's industry or the firm's city) are assumed predetermined.

The regressions in Table 4 test the relation with employee retention rates of three groups of variables: intra-firm remuneration equity variables (regression 1), intra-firm gender equity variables (regression 2), and variables related to the firm's industry and city (regression 3). The three groups of variables are combined in regression (4). All regressions include the lagged dependent variable, the variable "effective size" as an additional control—both are always significant—and year dummies (not shown).

In regression (1), the firm's average wage relative to the rest of its industry in the city and the dispersion of real wage adjustments are statistically significant (with 99.9 percent confidence) with the sign predicted by theory. The wage gap (90-50 percentile ratio) is also statistically significant, but not in the direction predicted by relative deprivation theory. In regression (2), the share of women and the interaction of this variable with the gender wage gap are highly significant with the expected signs. However, without the interaction, the gender wage gap is significant in the wrong direction (like the wage gap). In regression (3), consistently with theory, firms that operate in more concentrated industries in their city have higher employee retention rates, with high statistical significance. In this regression, employee retention is related directly and significantly with the nation-wide industry shocks and with the level of government expenditure per capita in the city. All the previous results remain when the three sets of explanatory variables are combined in regression (4). Excluding the only non-significant variable of regression (4) does not alter the significance of the remaining explanatory variables (results not shown). The statistical tests at the bottom of the table indicate that there is no evidence of model miss-specification or auto-correlation of the residuals

The parameters estimated in regression (4) can be used to calculate the changes in the employee retention rate associated with typical changes in the explanatory variables (Table 5). For each explanatory variable, the same-year change in the dependent variable is the product of the coefficient estimated in regression 4 of Table 4 and the standard deviation of the explanatory variable (Table 2). When the explanatory variable forms part of an interaction, the interaction effect is added, evaluated at the mean value of the interacting variable. The firm's average wage relative to the rest of the industry in the same city has the largest simulated impact on the employee retention rate of a typical firm: one standard deviation of the explanatory variable is associated with an increase of 11.4 percent points in the employee retention rate (which corresponds to 56 percent of the standard deviation of the employee retention rate across all firms). Second in magnitude is the dispersion of real wage adjustments within the firm: flattening real wage adjustments, one standard deviation is associated with a 7.3 percent point increase in retention, which is equivalent to 36 percent of its own standard deviation. Following in order by impact is the share of women: 4.1 percent points of additional employee retention, including the effect of its interaction with the gender wage gap. As the last column of Table 5 shows, long-term effects are approximately 16 percent larger than the same-year effects just mentioned, as higher retention rates reinforce themselves through time (assuming that the

change in the explanatory variable is permanent and other explanatory variables remain unchanged).

These results must be interpreted with caution because strict causality is not guaranteed and because the explanatory variables may be interdependent, even in the same year. For instance, raising the firm's average wage with respect to the rest of the industry in the city may result in changing the intra-firm wage gap, the average real wage adjustments, and the dispersion of real wage adjustments, all of which may impinge in workers' decisions to stay or quit.

It must also be kept in mind that the effects of the explanatory variables on retention rates may not be the same for firms of different sizes, located in cities of different size or that operate in different sectors (see Appendix tables for details). Variables that are not significant in the regression that uses the full sample of firms may be significant for specific groups. It is the case of the average real wage adjustment, which is not significant in the whole sample, but it is among some groups of firms (directly in some groups, inversely in others). Also, variables that are significant for the whole universe of firms may not be significant or may even change sign for some groups of firms. The latter is the case of the gender wage gap, which is directly associated with personnel retention in the universe of firms (at odds with theory) but is inversely associated in the largest firms and in group of firms that operate in the smallest cities.

Table 4. System dynamic panel data estimations of employee retention rates.

	Regression (1)	Regression (2)	Regression (3)	Regression (4)
Lagged dependent variable	0.139***	0.158***	0.158***	0.137***
	(0,0043)	(0,0042)	(0,0043)	(0,0042)
Firm level variables				
Effective size (logs)	-0.0342***	-0.0603***	-0.0744***	-0.0235***
	(0,0022)	(0,0025)	(0,0025)	(0,0022)
Average wage relative to rest of industry in city (logs)	0.252***			0.216***
	(0,0097)			(0,0093)
Wage gap (90-to-50 percentile ratio)	0.0755***			0.0624***
	(0,0068)			(0,0065)
Average real wage adjustments (logs)	0,00701			-0,0121
	(0,0142)			(0,0134)
Dispersion of real wage adjustments	-0.771***			-0.706***
	(0,0230)			(0,0221)
Share of women		0.162***		0.140***
		(0,0137)		(0,0135)
Gender diversity		0.0651***		0.0899***
		(0,0118)		(0,0109)
Gender wage gap (logs)		0,0128		0,00415
		(0,0078)		(0,0081)
Industry and city level variables				
Herfindahl concentration index of industry in city			0.250***	0.159***
			(0,0100)	(0,0091)
Bartik instrument of nation-wide industry shocks (logs)			0.0073***	0.0074***
			(0,0011)	(0,0010)
Government expenditure per capita (logs)			0.0219***	0.0182***
			(0,0038)	(0,0038)
Number of observations (firms in city in year)	334698	334698	334698	334698
Number of firms (firms in city)	61887	61887	61887	61887
Number of instruments in estimation	190	170	71	295
Sum of squared differenced residuals	8792,1	8188,8	8495,8	8313,6
Chi-squared	8484,3	4874,4	4573,1	9549,4
Arellano-Bond test for zero autocorrelation, order 1	-101.02***	-101.92***	-102.37***	-100.79***
Arellano-Bond test for zero autocorrelation, order 2	6.29***	4.78***	4.09***	6.31***

Regression (1) tests the relation between employee retention and intra-firm equity variables. Regression (2) considers gender equities. Regression (3) considers industry and city variables. The three sets of explanatory variables are combined in Regression (4).

Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All the regressions include a constant term and year dummies.

Source: own calculations from Ministry of Health's PILA 2008-2016.

Table 6 presents the estimated long-term effects of one standard deviation changes of the three firm-level explanatory variables with the most consistent impact predicted by theory on employee retention in the groups of firms considered. The computations are made with the same method of Table 5, with the coefficients and the standard deviations of the corresponding groups (Appendix tables A1, A2, and A3 show the group regressions from which the coefficients are taken). Retention rates are consistently associated (directly) with the firm's average wage with respect to the rest of the industry in the city, (inversely) with the intra-firm dispersion of real wage adjustments, and (directly) with the share of women (the calculations in this case consider the interaction effect, which is evaluated at the mean value of the interacting variable, namely the gender wage gap).

Table 5. Estimated impact on employee retention rates of one standard deviation of change in explanatory variables.

	Coefficients (from Regression 4 of Table 4)		Standard deviation of variable (from Tables 1 and 2)	Same year effect on employee retention rate (coefficient* standard deviation)	Long-term effect on employee retention rate (coefficient* standard deviation/(1-coefficient of lagged dependent variable))
Lagged dependent variable	0,1370	***	0,2022	...	...
Firm level variables					
Effective size (logs)	-0,0235	***	1,1458	-0,0269	-0,0312
Average wage relative to rest of industry in city (logs)	0,2160	***	0,4594	0,0992	0,1150
Wage gap (90-to-50 percentile ratio)	0,0624	***	0,4024	0,0251	0,0291
Average real wage adjustments (logs)	-0,0121		0,1046	-0,0013	-0,0015
Dispersion of real wage adjustments	-0,7060	***	0,0894	-0,0631	-0,0731
Share of women	0,1040	***	0,2378	0,0247	0,0287
Gender diversity	0,0899	***	0,2529	0,0227	0,0263
Gender wage gap (logs)	0,0042		0,2385	0,0010	0,0011
Industry and city level variables					
Herfindahl concentration index of industry and city	0,1590	***	0,2130	0,0339	0,0393
Bartik instrument of nation-wide industry shocks (logs)	0,0074	***	0,2748	0,0020	0,0024
Government expenditure per capita (logs)	0,0182	***	0,3662	0,0067	0,0077

Source: own calculations from Ministry of Health's PILA 2008-2016. * p<0.05, ** p<0.01, *** p<0.001.

Table 6. Estimated same-year effect on employee retention of one standard deviation change of selected equity variables by groups of firm size, city size and industry quartiles of complexity

	Average wage relative to rest of industry in city	Intra-firm dispersion of real wage adjustments	Share of women
All urban firms with at least 10 employees	0,100	-0,067	0,036
By firm size (by effective number of employees):			
10 to 25 employees	0,076	-0,074	0,041
25 to 50 employees	0,090	-0,052	0,026
50 to 100 employees	0,109	-0,042	0,022
100 or more employees	0,123	-0,034	0,014
By city size (by working age population):			
Largest 6 cities	0,093	-0,064	0,035
Medium-large 15 cities	0,103	-0,061	0,022
Medium-small 16 cities	0,081	-0,060	0,038
Smallest 25 cities	0,040	-0,056	0,064
By industry quartiles of complexity:			
Lowest quartile (110 industries)	0,093	-0,063	0,038
Second quartile (110 industries)	0,117	-0,052	0,030
Third quartile (110 industries)	0,058	-0,057	0,019
Highest quartile (110 industries)	0,086	-0,054	...

Method: the effects are computed as the product of the coefficients estimated (from regressions in Tables 4, A1, A2, and A3) and the standard deviations of the corresponding variables by group.
... indicates that the coefficient is not statistically significant with at least 95% confidence.
Source: own calculations from Ministry of Health's PILA 2008-2016.

6. Summary and discussion

This paper seeks to advance our understanding of how equity considerations influence the decision of workers to stay or leave their jobs. Although previous empirical literature has paid attention to many determinants of personnel retention, a comprehensive analysis of how equity variables affect turnover is expected. Also, although abundant studies have focused on specific firms, industries and/or professions, only a few studies (for developed countries) have taken advantage of the growing availability of matched employer-employee panel data that provide information on all registered firms.

This paper presented evidence for the universe of Colombian urban firms with at least 10 employees during 2008-2016, and the relation between intra-firm equity and employee retention. Colombia is a fitting case for this topic because, in international comparisons, labour earnings are very concentrated, salaried workers change jobs with high frequency, and most firms are poorly managed. Based on equity theory, efficiency wage theory, relative deprivation theory, and gender theory, a set of explanatory variables was chosen to study the association between intra-firm equity and employee retention. The Blundell and Bond (1998) method of estimation was adopted, which is suitable for panel data with large cross-sectional variation but few years of time variation when the error term and some of the explanatory variables may be endogenous to the dependent variable.

The estimations gave strong support to the main hypotheses derived from equity, efficiency wage, and diversity theories, but not to a hypothesis derived from relative deprivation theory, i.e., employee retention is lower when intra-firm wage gaps are larger (the opposite was found consistently). Apart from the latter, the equity variables most strongly and consistently associated with employee retention in the whole universe of firms and across groups of firms (of different size, operating in cities of different sizes, or in different sectors) were firm's average wages relative to the rest of the industry in the city of operation of the firm (directly), the intra-firm dispersion of real wage adjustments (inversely), and the share of women (directly), except in the firms operating in the smallest cities and in finance, utilities, and transportation.

This study also uncovered evidence of the relationship between employee retention and some other variables. Retention rates are higher in smaller firms, in those operating in more locally concentrated industries, in industries facing favourable nation-wide shocks, and in cities that have higher government expenditure per capita. However, these relationships do not hold consistently across groups, thus indicating the presence of heterogeneities.

Simple calculations using the estimated coefficients and the standard deviations of the explanatory variables suggest that equity considerations play an important role in employees' decisions to stay or quit their jobs. Although this study makes no claims of causality, the results suggest that raising average wages with respect to those paid by other firms in the same industry and city, flattening intra-firm wage adjustments and increasing female participation, may substantially improve the ability of firms to retain their employees.

Some of the results of this study are at odds with human resource management principles advocated by some organisations. According to the World Management Survey, good human resource management implies remunerating employees according to their performance, and not simply based on tenure, as most Colombian firms do (Bloom et al., 2012). But our results strongly suggest that flat wage adjustments are conducive to higher employee retention rates. However, our finding that employee retention is higher in firms with larger wage gaps between top and middle-level workers—which is at odds with relative deprivation theory—does not contradict human resource management principles because a strong management hierarchy, with highly paid employees in the upper echelons of the firm, may be a necessary condition for the efficient performance of the whole employee team, especially if the firm is large and highly sophisticated.

Based on the World Management Survey, Bloom et al. (2012) have argued that cultural differences may be behind some of the stark differences in human resource management practices across countries. The results of this paper may shed some light on the issue, especially regarding the role of gender. According to the results, more female participation is associated with higher employee retention in the universe of firms, but not in some groups of firms. The main source of heterogeneity is the size of the city: in the smallest cities, where more traditional values about the role of women tend to prevail, higher female participation is (weakly) associated with lower retention.

However, culture may be only one among many factors influencing equity considerations and their effect on employee retention. If enhancing intra-firm equity entails pecuniary costs that override the potential productivity gains of having more motivated and stable employees, low wages, uneven wage adjustments, and discriminatory treatment of women may be rational strategies from the point of view of profit maximising firms. Nevertheless, such strategies may generate negative externalities if they deteriorate workers' wellbeing and if, as mentioned in the introduction, firms choose not to invest in training, thus tipping the economy into a low productivity trap. These issues are fertile ground for future research.

Conflict of interest

The authors declare no conflict of interest.

References

Arellano, Manuel, and Stephen Bond. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies 58*(2), 277-297.

Backes-Gellner U, D. Bessey, K. Pull K and S. Tuor. (2008). "What behavioural economics teaches personnel economics." *Die Unternehmung 62*, 217–234.

Barth, Erling, and Harald Dale-Olsen. (2009). "Monopsonistic discrimination, worker turnover, and the gender wage gap." *Labour Economics 16*(5), 589-597.

Berger, J. E., and R. L. Boyle Jr. (1992). "How to avoid the high costs of physician turnover." *Medical Group Management Journal 39*(6), 80-82.

Bergman, Rebecca, Shoshana Eckerling, Hava Golander, Rena Sharon, and Adrian Tomer. (1984). "Staff composition, job perceptions, and work retention of nursing personnel in geriatric institutions." *International Journal of Nursing Studies 21*(4), 279-293.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen. (2012). "Management practices across firms and countries." *Academy of Management Perspectives 26*(1), 12-33.

Blundell, Richard, and Stephen Bond. (1998). "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics 87*(1), 115-143.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. (2018)."The morale effects of pay inequality." *The Quarterly Journal of Economics 133*(2), 611-663.

Burdett, Kenneth, and Dale T. Mortensen. (1998). "Wage differentials, employer size, and unemployment." *International Economic Review*, 257-273.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. (2012). "Inequality at work: The effect of peer salaries on job satisfaction." *American Economic Review 102*(6), 2981-3003.

Chen, Hsing-Chu, Cheng-I. Chu, Ying-Hui Wang, and Ling-Chu Lin. (2008). "Turnover factors revisited: a longitudinal study of Taiwan-based staff nurses." *International Journal of Nursing Studies 45*(2), 277-285.

Cowherd, Douglas M., and David I. Levine. (1992). "Product quality and pay equity between lower-level employees and top management: An investigation of distributive justice theory." *Administrative Science Quarterly*, 302-320.

Crosby, Faye. "Relative deprivation in organizational settings. (1984). "In Barry M. Staw and L.L. Cummings (eds.), *Research in Organizational Behavior, 6*, 51-93.

Das, Bidisha Lahkar, and Mukulesh Baruah. (2013). "Employee retention: A review of literature." *Journal of Business and Management 14*(2), 8-16.

Dohmen T. (2014). Behavioral labor economics: Advances and future directions. *Labour Economics 30*, 71–85.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. (2009). "Homo reciprocans: Survey Evidence on Behavioural Outcomes." *The Economic Journal 119*(536), 592-612

Doig, Kathy, and Susan Beck. (2005). "Factors contributing to the retention of clinical laboratory personnel." *American Society for Clinical Laboratory Science 18*(1) 16-27.

Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard. (2019). "Fairness and frictions: The impact of unequal raises on quit behavior." *American Economic Review 109*(2): 620-63.

Ely, Robin J., and Debra E. Meyerson. (2000). "Theories of gender in organizations: A new approach to organizational analysis and change." *Research in Organizational Behavior 22*, 103-151.

Faris, John H. (1984). "Economic and noneconomic factors of personnel recruitment and retention in the AVF." *Armed Forces & Society 10*(2), 251-275.

Flint, Douglas, Lynn M. Haley, and Jeffrey J. McNally. (2013). "Individual and organizational determinants of turnover intent." *Personnel Review*.

Frederiksen, Anders. (2017). "Job Satisfaction and Employee Turnover: A firm-level perspective." *German Journal of Human Resource Management 31*(2), 132-161.

Gilroy, Curtis L., David K. Horne, and D. Alton Smith. (1991). Military Compensation and Personnel Retention. Models and Evidence. *Army Research Institute for the Behavioral and Social Sciences*.

Glance, Natalie S., Tad Hogg, and Bernardo A. Huberman. (1997). "Training and turnover in the evolution of organizations." *Organization Science 8*(1), 84-96.

Grund, Christian, Alex Bryson, Robert Dur, Christine Harbring, Alexander K. Koch, and Edward P. Lazear. (2017). "Personnel economics: A research field comes of age." *German Journal of Human Resource Management 31*(2), 101-107.

Homans, George C. (1974). Social Behavior: Its Elementary Forms, rev. ed. New York: Harcourt, Brace, Jovanovich..

Inter-American Development Bank. (2020, October 20). Database: Labor Markets and Social Security Information System: https://www.iadb.org/en/sector/social-investment/sims/home.

Jha, Shweta. (2009). "Determinants of employee turnover intentions: A review." *Management Today, 9*(2).

Koster, Fleur, Andries De Grip, and Didier Fouarge. (2011). "Does perceived support in employee development affect personnel turnover?" The *International Journal of Human Resource Management, 22*(11), 2403-2418.

Kremer, Michael. (1993). "The O-ring theory of economic development." *The Quarterly Journal of Economics 108*(3): 551-575.

Manning, Alan. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Martin, Joanne. (1981). "Relative deprivation: A theory of distributive injustice for an era of shrinking resources." In L. L. Cummings and Barry M. Staw (eds.), *Research in Organizational Behavior, 3*, 53-107.

Mohrenweiser, J. and Pfeifer, C. (2019). Firms' Wage Structures, Workers' Fairness Perceptions, Job Satisfaction and Turnover Intentions: Evidence from Linked Employer-Employee Data. *IZA Discussion Paper Series IZA DP 12821*.

Nielsen, Vibeke Lehmann, and Mikkel Bo Madsen. (2017). "Does Gender Diversity in the Workplace Affect Job Satisfaction and Turnover Intentions? " *International Public Management Review 18*(1), 77-115.

O'Connell, Matthew, and Mei-Chuan Kung. (2007). "The Cost of Employee Turnover." *Industrial Management 49*(1).

Orrick, Dwayne W. (2008). Recruitment, retention, and turnover of police personnel: Reliable, practical, and effective solutions. Charles C Thomas Publisher.

Pfeifer, Christian, and Stefan Schneck. (2012). "Relative wage positions and quit behavior: Evidence from linked employer-employee data." *ILR Review 65*(1), 126-147.

Ramlall, Sunil. (2004). "A review of employee motivation theories and their implications for employee retention within organizations." *Journal of American Academy of Business 5*(1/2), 52-63.

Rodríguez-Castelán, Carlos, Luis F. López-Calva, Nora Lustig, and Daniel Valderrama. (2016). Understanding the dynamics of labor income inequality in Latin America. The World Bank.

Schlicht, Ekkehart. (1978). "Labour turnover, wage structure, and natural unemployment." Zeitschrift für die Gesamte Staatswissenschaft/ *Journal of Institutional and Theoretical Economics*, 337-346.

Scur, Daniela, Raffaella Sadun, John Van Reenen, Renata Lemos, and Nicholas Bloom. (2021). "The World Management Survey at 18: lessons and the way forward." *Oxford Review of Economic Policy 37*(2), 231-258.

Steel, Robert P., and Nestor K. Ovalle. (1984). "A review and meta-analysis of research on the relationship between behavioral intentions and employee turnover." *Journal of Applied Psychology 69*(4), 673.

Stiglitz, Joseph E. (1974). "Alternative theories of wage determination and unemployment in LDC's: The labor turnover model." *The Quarterly Journal of Economics 88*(2), 194-227.

Walster, Elaine, G. William Walster, and Ellen Berscheid. (1978). Equity: Theory and Research. Boston: Allyn and Bacon.

Woolley, Anita Williams, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. (2010). "Evidence for a collective intelligence factor in the performance of human groups." *Science 330*(6004), 686-688

Annexes

Table A1. System dynamic panel data estimations of employee retention rates by group of firm size.

	10 to 25 employees	25 to 50 employees	50 to 100 employees	100 or more employees
Lagged dependent variable	0.108***	0.164***	0.211***	0.273***
	(0,0055)	(0,0089)	(0,0123)	(0,0131)
Firm level variables				
Effective size (logs)	-0.0496***	-0.0189***	-0,00579	0.0177***
	(0,0029)	(0,0031)	(0,0038)	(0,0035)
Average wage relative to rest of industry in city (logs)	0.166***	0.197***	0.238***	0.267***
	(0,0146)	(0,0165)	(0,0213)	(0,0152)
Wage gap (90-to-50 percentile ratio)	0.0585***	0.0970***	0.0693***	0.102***
	(0,0084)	(0,0135)	(0,0172)	(0,0164)
Average real wage adjustments (logs)	-0.0665**	-0,024	0,0028	0.0845***
	(0,0231)	(0,0263)	(0,0234)	(0,0193)
Dispersion of real wage adjustments	-0.762***	-0.622***	-0.526***	-0.488***
	(0,0345)	(0,0460)	(0,0423)	(0,0409)
Share of women	0.170***	0.109***	0.0967**	0.0600*
	(0,0195)	(0,0252)	(0,0294)	(0,0266)
Gender diversity	0.122***	0.103***	0,0123	-0.134***
	(0,0155)	(0,0193)	(0,0244)	(0,0207)
Gender wage gap (logs)	0,00826	-0,00311	0,0183	-0,0358
	(0,0103)	(0,0168)	(0,0221)	(0,0242)
Industry and city level variables				
Herfindahl concentration index of industry in city	0.0997***	0.139***	0.189***	0.135***
	(0,0127)	(0,0181)	(0,0241)	(0,0220)
Bartik instrument of nation-wide industry shocks (logs)	0.00567**	0.00505*	0.00970***	0.0149***
	(0,0017)	(0,0022)	(0,0025)	(0,0022)
Government expenditure per capita (logs)	0.0254***	0,00281	0.0208*	0.0227*
	(0,0059)	(0,0076)	(0,0094)	(0,0090)
Number of observations (firms in city in year)	175316	73625	41747	44010
Number of firms (firms in city)	34327	13347	7168	7045
Number of instruments in estimation	295	295	295	295
Sum of squared differenced residuals	5645	1604,9	725,5	578,7
Chi-squared	4274	2535,4	1950,5	2948,1
Arellano-Bond test for zero autocorrelation, order 1	-72.70***	-47.32***	-33.71***	-31.88***
Arellano-Bond test for zero autocorrelation, order 2	4.01***	3.30***	1,02	5.01***

Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All the regressions include a constant term and year dummies.
Source: own calculations from Ministry of Health's PILA 2008-2016.

Table A2. System dynamic panel data estimations of employee retention rates by group of city size.

	Largest 6 cities	Medium-large 15 cities	Medium-small 16 cities	Smallest 25 cities
Lagged dependent variable	0.131***	0.157***	0.151***	0.150***
	(0,0048)	(0,0093)	(0,0184)	(0,0256)
Firm level variables				
Effective size (logs)	-0.0274***	-0.00929*	-0,00625	-0,00167
	(0,0028)	(0,0038)	(0,0061)	(0,0102)
Average wage relative to rest of industry in city (logs)	0.195***	0.247***	0.203***	0.115**
	(0,0107)	(0,0182)	(0,0284)	(0,0415)
Wage gap (90-to-50 percentile ratio)	0.0613***	0.0692***	0.0756**	0.102***
	(0,0071)	(0,0153)	(0,0288)	(0,0309)
Average real wage adjustments (logs)	-0.0417*	0,0251	-0,0468	-0.0746*
	(0,0194)	(0,0202)	(0,0252)	(0,0374)
Dispersion of real wage adjustments	-0.714***	-0.705***	-0.629***	-0.586***
	(0,0281)	(0,0397)	(0,0689)	(0,0747)
Share of women	0.147***	0.0934***	0.155***	0.258***
	(0,0165)	(0,0249)	(0,0435)	(0,0615)
Gender diversity	0.0912***	0.0844***	0,0351	0,0441
	(0,0131)	(0,0214)	(0,0406)	(0,0504)
Gender wage gap (logs)	0,0117	-0,0329	-0,00939	-0,062
	(0,0089)	(0,0204)	(0,0381)	(0,0418)
Industry and city level variables				
Herfindahl concentration index of industry in city	0.195***	0.143***	0.110***	0.143***
	(0,0147)	(0,0144)	(0,0216)	(0,0337)
Bartik instrument of nation-wide industry shocks (logs)	0.00344**	0.0179***	0.0241***	0,0131
	(0,0012)	(0,0024)	(0,0053)	(0,0077)
Government expenditure per capita (logs)	0.0215***	0,00752	-0.0515***	0,0164
	(0,0056)	(0,0082)	(0,0136)	(0,0112)
Number of observations (firms in city in year)	239126	69217	17108	9247
Number of firms (firms in city)	43403	13168	3412	1904
Number of instruments in estimation	295	295	295	295
Sum of squared differenced residuals	6218,5	1486,8	380,9	201,8
Chi-squared	6583,6	2542,3	659,8	292,5
Arellano-Bond test for zero autocorrelation, order 1	-85.62***	-43.83***	-22.79***	-16.12***
Arellano-Bond test for zero autocorrelation, order 2	4.71***	4.48***	1.18	1.88

Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All the regressions include a constant term and year dummies.
Source: own calculations from Ministry of Health's PILA 2008-2016.

Table A3. System dynamic panel data estimations of employee retention rates by group of industry complexity.

	Lowest quartile of industry complexity	Second quartile of industry complexity	Third quartile of industry complexity	Highest quartile of industry complexity
Lagged dependent variable	0.130***	0.139***	0.165***	0.143***
	(0,0065)	(0,0095)	(0,0110)	(0,0145)
Firm level variables				
Effective size (logs)	-0.0243***	-0,00727	-0.0172***	-0,00746
	(0,0035)	(0,0045)	(0,0048)	(0,0062)
Average wage relative to rest of industry in city (logs)	0.210***	0.249***	0.125***	0.184***
	(0,0140)	(0,0169)	(0,0170)	(0,0205)
Wage gap (90-to-50 percentile ratio)	0.0621***	0.0618***	0.0476***	0,0278
	(0,0103)	(0,0133)	(0,0138)	(0,0167)
Average real wage adjustments (logs)	-0,00314	0,00298	-0.0636*	-0,0303
	(0,0228)	(0,0189)	(0,0273)	(0,0333)
Dispersion of real wage adjustments	-0.683***	-0.599***	-0.699***	-0.675***
	(0,0365)	(0,0450)	(0,0479)	(0,0629)
Share of women	0.158***	0.126***	0.0831**	0,0663
	(0,0210)	(0,0298)	(0,0279)	(0,0413)
Gender diversity	0.0875***	0.0521*	0,0389	0.0742*
	(0,0172)	(0,0237)	(0,0261)	(0,0345)
Gender wage gap (logs)	0,0101	-0,00838	0,0133	0,00444
	(0,0134)	(0,0176)	(0,0174)	(0,0223)
Industry and city level variables				
Herfindahl concentration index of industry in city	0.189***	0.0591***	0.0447*	0,013
	(0,0182)	(0,0165)	(0,0192)	(0,0179)
Bartik instrument of nation-wide industry shocks (logs)	0.00429*	0.00708***	0.00788**	0.0119***
	(0,0020)	(0,0019)	(0,0029)	(0,0024)
Government expenditure per capita (logs)	0.0202**	0.0352***	0.0267**	0,00923
	(0,0066)	(0,0086)	(0,0086)	(0,0103)
Number of observations (firms in city in year)	122845	61494	50648	32023
Number of firms (firms in city)	23838	10996	8101	5109
Number of instruments in estimation	295	295	295	295
Sum of squared differenced residuals	3225,1	1353,8	1031	610
Chi-squared	3540,6	1910,7	1465,1	834,7
Arellano-Bond test for zero autocorrelation, order 1	-62.19***	-44.25***	-39.57***	-30.28***
Arellano-Bond test for zero autocorrelation, order 2	3.46***	4.05***	3.98***	0.56

Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All the regressions include a constant term and year dummies.
Source: own calculations from Ministry of Health's PILA 2008-2016.