Skill mismatch and labour turnover in a developing country: the Colombian case

By: Luz Adriana Flórez
Leidy Gómez D.

Borradores de ECONOMÍA

No. 1099
2019

Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia

Skill mismatch and labour turnover in a developing country: the Colombian case

Luz Adriana Flórez♣
lflorefl@banrep.gov.co

Leidy Gómez D.♠
lejgomezdu@gmail.com

Banco de la República

The opinions expressed here are not necessarily those of neither the Banco de la República nor its Board of Directors

Abstract

The objective of this paper is to analyze the impact of *skill mismatch* on labour turnover for the case of Colombia. Our work follows the of the job matching theory of Jovanovic (1979a, 1979b, 1984). In line with this theory we find a positive relationship between *skill mismatch* and labour turnover (measured as the worker reallocation rate) using a panel of 23 cities for the period 2009-2017. Our results suggest that cities with a higher proportion of mismatched workers present higher worker reallocation rates. In this case one standard deviation of increment in the proportion of *mismatch* workers increases the WR rate around 0.12 standard deviations. This result is explained mainly by the increase on separations as is suggested by the theory.

JEL classification: I25, J62, J63, J64.

Keywords: Skill mismatch, overqualification, underqualification, worker reallocation, panel data, and cross-sectional dependence.

♣ Researcher at the Central Bank of Colombia (Banco de la República, Medellín branch) and ♠ research assistant from National University (Medellín branch). The opinions expressed here are those of the authors and not of the *Banco de la República de Colombia* nor of its Board. We thank Andres Garcia, Luis Eduardo Arango, Blendi Kajsiu and the participants of the 5th Seminar of Applied Economy at Medellin University for their useful comments on an earlier version of this paper. All errors are authors´ responsibility. Corresponding author: lflorefl@banrep.gov.co.

Desajuste de habilidades y rotación laboral en un país en desarrollo: el caso colombiano

Luz Adriana Flórez♣
lflorefl@banrep.gov.co

Leidy Gómez D.♠
lejgomezdu@gmail.com

Banco de la República

Las opiniones contenidas en el presente documento son responsabilidad exclusiva de los autores y no comprometen al Banco de la República ni a su Junta Directiva.

Resumen

El objetivo de este documento es analizar el efecto del desajuste en habilidades en la rotación laboral para el caso de Colombia. Nuestro enfoque teórico sigue de cerca el modelo de búsqueda de empleo propuesto por Jovanovic (1979a, 1979b, 1984). Como es sugerido por esta literatura, usando el panel de 23 ciudades para el periodo 2009-2017, encontramos evidencia de una relación positiva entre el nivel del desajuste en las habilidades y la rotación laboral, medida como la reasignación de trabajadores. Un incremento de una desviación estándar en el nivel de desajuste de habilidades produce un incremento de 0.12 desviaciones estándar en la rotación laboral. Estos resultados se explican principalmente por el incremento en las separaciones como es sugerido por el modelo teórico.

Clasificación JEL: I25, J62, J63, J64.

Palabras claves: Desajuste de habilidades, sobre-calificación, baja calificación, datos panel, reasignación de trabajadores, y dependencia transversal.

♣ Investigadora del Banco Central de Colombia (Banco de la República, Sucursal Medellín) y ♠ asistente de investigación de la Universidad Nacional (Sede Medellín). Las opiniones expresadas aquí son de los autores y no comprometen al *Banco de la República de Colombia* ni a su Junta Directiva. Agradecemos los comentarios de Andres Garcia, Luis Eduardo Arango, Blendi Kajsiu y de los participantes del 5th Seminario de Economía Aplicada de la Universidad de Medellín. Todos los errores son responsabilidad de los autores. Autor corresponsal: lflorefl@banrep.gov.co.

1. Introduction

In the literature *skill mismatch* is defined, for a particular match, as the difference between the level of skills demanded by the firms and the skills offered by the workers. Hence, optimal allocation in the labour market exists when workers end up in jobs for which they were trained and which require the skills that they possess (van der Velden & Verhaest, 2017). In general, the evidence shows that *skill mismatch* is costly for both firms and workers.

The literature, mainly for developed economies, has shown that those individuals who are in a *mismatch* situation have lower job satisfaction, lower tenure, lower wages, and are more likely to look for new jobs (Alba-Ramirez, 1993; Allen & Van der Velden, 2001; Badillo-Amador & Vila, 2013; Di Pietro & Urwin, 2006; Verhaest & Omey, 2006). In the case of the firm, a *mismatch* can decrease its level of productivity and increase labour turnover (Hersch, 1991; McGowan & Andrews, 2017; Verhaest & Omey, 2006). This negative effects in human capital and the level of productivity, can be reflected in a high unemployment equilibrium and low GDP growth (Lucifora & Origo, 2002; Manacorda & Petrongolo, 1999; Olitsky, 2008; Quintini, 2011).

The study of *skill mismatch*[1] is particularly relevant in the case of Colombia due to its high level compared with other developed and developing countries. **Figure 1** presents the level of *mismatch* divided by overqualified and underqualified workers. A worker is overqualified when his/her level of education exceeds his/her job requirement, and underqualified in the contrary case. In Colombia the level of *skill mismatch* is twice the average of the OECD countries, and more than twice the one observed in the USA. Moreover, the *skill mismatch* in Colombia is also higher compared with similar Latin American countries such as Perú, Chile, Brazil, Argentina and México.

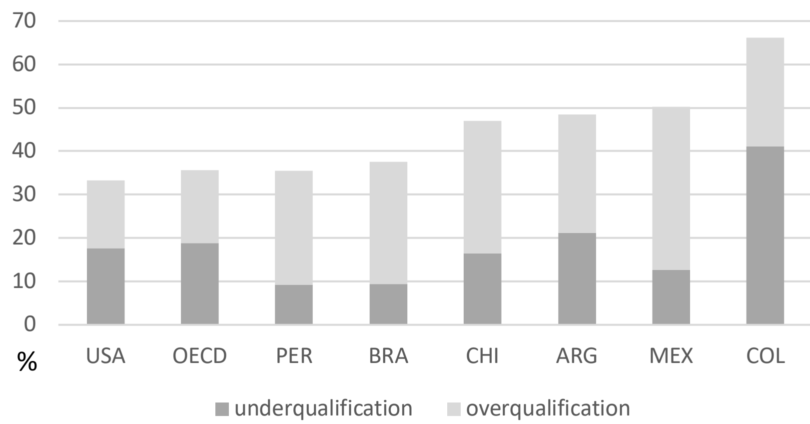
The objective of this paper is to analyze the effect of *skill mismatch* on the Colombian labour turnover, following the theoretical guidance of the job matching theory (Jovanovic, 1979a, 1979b, 1984). The model assumes that there is imperfect information, so the quality of the match between a worker and a firm is unknown at the beginning, but reveals itself as the

[1] To be able to compared the Colombian skill-mismatch measured with the one used in the OECD, we use the "*skill mismatch-national*" definition. This measure is defined with more details in section 4.

match produces output over time. Therefore, workers and firms update their beliefs on the quality of the match in two ways: through firms screening and through the observation of the output. If the current match is a disappointment for the worker, and new offers do not arrive, he may choose to be unemployed. Then, this model predicts that workers who are in a *mismatch* situation have a higher labour turnover.

Figure 1: *Skill mismatch* in developed and developing countries

(percentage of total workers-2018)



Source: OECD (2017), Skills for Jobs Database and author's calculations based on the Colombian household survey- GEIH. Note: A worker is overqualified if his/her level of education exceeds his job requirements and underqualified in the contrary case.

The evidence on the relationship between *skill mismatch* and labour turnover comes mainly from developed countries. Researchers have found that *mismatched* workers have a higher probability of leaving employment than matched workers (Hersch, 1995; McGuinness & Wooden, 2007; Robst, 1995; Rubb, 2013; Verhaest & Omey, 2006), have less experience or tenure (Alba-Ramirez, 1993; Alba-Ramírez & Blázquez, 2003; Sloane, Battu, & Seaman, 1999) and are involved in job search while employed (Di Pietro & Urwin, 2006; Groot & Maassen van den Brink, 2003). However, empirical findings for developing economies are scarce.

The main contribution of this paper is to provide empirical evidence on the relationship between *skill mismatch* and labour reallocation in a developing country such as Colombia, as suggested by the job matching theory (Jovanovic, 1979a, 1979b, 1984). To do so, we use data from the Integrated Record of Contributions to Social Security (PILA by its Spanish acronym) and the official Colombian Household Survey (GEIH by its initials in Spanish) to build an aggregate panel data at the city level (23 main cities in Colombia)[2], with a quarterly frequency for the period 2009-2017. Our results indicate, that cities with a high *skill mismatch* present a high labour reallocation (measured as the worker reallocation rate), in line with our theoretical framework. In this case one standard deviation of increment in the proportion of *mismatch* workers increases the WR rate around 0.12 standard deviations. This result is explained mainly by the increase on separations as is suggested by the theory. Our results hold even we control for the productive structure or specific labour demand of each city or for specific institutions such as the minimum wage.

The rest of the paper is organized as follows. The next section presents a literature review on the effects of *skill mismatch*, focusing mainly on its effects on labour turnover. Section 3 presents the theoretical model (Jovanovic, 1979a, 1979b, 1984) that informs our empirical strategy. Section 4 describes the data used in this study and section 5 discusses our empirical strategy and reports our main estimates, including some robustness checks. Finally, section 6 summarize the main findings and concludes.

2. Literature review

Empirical evidence about the effect of *skill mismatch* is extensive in developed economies and can be classified into two groups. The first approach studies the wage returns and level of productivity of those who are in a *mismatch*, while the second one tries to understand the existence of *mismatch*. In the second approach, we can find two models commonly used to explain the reason for the *mismatch*: the occupational mobility theory and job matching theory (Jovanovic, 1979a, 1979b, 1984).

The first empirical approach has shown that overqualified workers face wage penalties. This implies that they earn less than those who have the same educational level and are employed in a job that adequately fits their skills. On the contrary, underqualified workers present higher wages than those with the same educational level, but in jobs for which they are adequately educated. (See Allen & Van der Velden, 2001; Badillo-Amador & Vila, 2013;

[2] The main twenty-three cities are: Bogotá, Medellín and its metropolitan area, Cali and its metropolitan area, Barranquilla and its metropolitan area, Bucaramanga and its metropolitan area, Pasto, Cartagena, Cúcuta and its metropolitan area, Neiva, Pereira and its metropolitan area, Montería, Villavicencio, Tunja, Quibdó, Popayán, Ibagué, Valledupar, Sincelejo, Riohacha, Florencia, Santa Marta and Armenia.

Dolton & Vignoles, 2000; Mcguinness, 2006; Rumberger, 1987; Verdugo & Verdugo, 1989). Hartog (2000) studies five countries (Netherlands, Spain, Portugal, United Kingdom and United States), to show that wage returns for overeducated workers are typically about half to two-thirds of the returns to those who are adequately educated.

In addition to the negative wage returns of those overqualified workers, effects on productivity have been also studied; although here the results are less conclusive. McGowan and Andrews (2017) argue that *mismatch* is associated with low labour productivity and conclude that an increase of one standard deviation in the overqualification measure, is associated with a 4% reduction in the overall labour productivity level. In contrast, Kampelmann and Rycx (2012), using as measure of labour productivity by firm the added value per worker, show that additional years of overeducation are positive to the labour productivity of the firm, while additional years of undereducation are detrimental.

The second approach focuses on understanding the reason why the *mismatch* phenomenon occurs despite its negative returns.[3] The occupational mobility theory (Rosen, 1972; Sicherman & Galor, 1990; Sicherman, 1991) explains the phenomenon of *mismatch* as temporary, where workers are willing to accept a *skill mismatch* with the expectation of a job promotion. Therefore, the occupational mobility theory states that overqualified workers have a higher labour turnover rate, since they are willing to accept a job for which they are overqualified with the expectation of gaining experience and obtaining a better job and wages in the future.

For the case of Spain, Alba-Ramirez (1993)[4] found that overqualified workers are younger, less experienced, more educated and have shorter on-the-job training activities than those adequately educated. Moreover, the authors found that these workers have a 4.7 times lower probability of never having changed jobs than those adequately educated, evidence that supports the hypothesis of higher labour mobility among those overqualified (similar results

[3] The returns on wages of those individuals who are in *mismatch* contradict the prediction of the human capital theory (Becker, 1964; Mincer, 1974), which predict that higher education levels imply higher productivity and, therefore, higher wages.

[4] These authors also find that underqualified workers are more likely to have more experience, less education, a full-time job and report more on-the-job training needed for the job.

were found by Frei and Sousa-Poza (2012)[5] using the Swiss Household Panel-SHP and Verhaest and Omey (2006) using a Flemish database). Contrary to these findings, there are some authors that do not find empirical evidence on the relation between *mismatch* and labour mobility (Serrano and Malo (1996) for the case of Spain; Wen and Maani (2019) for Australia, and Büchel and Mertens (2004) for Germany).

Another theory used to explain the cause of *skill mismatch* is the theory of job matching (Jovanovic, 1979a, 1979b, 1984). It assumes that given the imperfect information, only through the experience in the job, firms and workers can realize if there is a *mismatch* or not. The overqualified workers will try to get a better job position in accordance with their educational level, increasing their probability of voluntary leaving their current job. On the other hand, the underqualified ones, in spite of not having any incentive to quit, would be more likely to be fired by the firm. Then, the job matching model suggests that *skill mismatch* might cause a high labour turnover and separations. Following this theoretical model, several authors have explored the effects of the *skill mismatch* on different measures of the labour turnover: showing that those who are in a *mismatch* situation are more likely to move to another job (McCall, 1990; Miller, 1984; Rubb, 2013); have a high quit intention (Hersch, 1991, 1995; McGuinness & Wooden, 2007; Robst, 1995; Verhaest & Omey, 2006); have a low tenure (Alba-Ramirez, 1993; Alba-Ramírez & Blázquez, 2003; Sloane et al., 1999); and have a high probability of actively looking for another job (Di Pietro & Urwin, 2006; Groot & Maassen van den Brink, 2003).

Evidence for the case of developing countries is very limited. In the Colombian case, studies concerning *skill mismatch* are focused on the returns. Just like in the case of developed countries, Castillo (2007) and Mora (2008) show that in Colombia the returns for overqualified workers are lower than for those adequately educated, while the evidence for underqualified workers is less clear. Moreover, overqualified workers tend to be younger and living in larger cities (Quejada & Ávila, 2017). Domínguez (2009), on the other hand, finds a positive relationship between overqualification and the level of tightness of the formal sector. He argues that a low vacancy jobs and high competition in quality positions, allow

[5] However, these authors suggest that workers manage to improve their *match* also through a better assignment of task, more appropriate to their education level in their current job.

employers to demand higher levels of education. However, Herrera-Idárraga et al. (2013), find opposite results. Finally, Mora (2008) finds evidence of a positive relation between overqualification and labour mobility. The author concludes that overqualified workers are less likely to remain in the same occupation.

The present paper contributes to the previous literature for the case of Colombia, providing more evidence on the relation between *skill mismatch* and labour reallocation, following the job matching theory of Jovanovic (1979a, 1979b, 1984). We use information from PILA and GEIH to build an aggregate panel data at the city level (23 main cities in Colombia), with a quarterly frequency for the period 2009-2017. We find that one standard deviation of increment in the proportion of *mismatch* workers increases the WR rate around 0.12 standard deviations. Moreover, this result is explained mainly by the increase on separation as is suggested by the theory.

3. Theoretical framework

Our empirical analysis is based on the job matching theory proposed by Jovanovic (1979a, 1979b, 1984). In this section, we present a summarized version of his model (Jovanovic, 1984). The model assumes a production function with constant returns to scale and labour as the only input used in the production. The quality of the match between the worker and his employer is denoted by μ, and the average output produced over a period of time is denoted by $X(t)$[6]. Workers and firms are risk neutral, live infinitely and they discount the future at a common constant discount rate, r. The quality of the match is unknown at the beginning but reveals itself as the match produces output over time. Among all possible matches between workers and firms, it is assumed that the quality of the match has a normal distribution such as: $\mu \sim N(\bar{\mu}, \sigma_\mu^2)$.

It is assumed that workers and firms update their beliefs of the quality of the match in two ways:

i) At the start of the match, firms screen workers. The outcome of the screen is a "noisy" estimate of μ, $m = \mu + \epsilon$, where $\epsilon \sim N(0, \sigma_m^2)$. The smaller σ_m^2, the better

[6] The model define $X(t)$ as the worker´s cumulative output over a period of length t, then $X(t) = \mu t + \sigma_x Z(t)$, where $\sigma_x > 0$ is a known constant, identical over all possible matches and $Z(t)$ is the standard Wiener process.

the screen. If $\sigma_m^2 = \infty$, a match is a pure "experience good", nothing can be learned about the match before it produces output. But if $\sigma_m^2 = 0$, the match is a pure "inspection good", one can learn its quality merely by inspecting it.

ii) Once the match start producing, firms have an observation of the output of the match. Then, the match with screen outcome m, job tenure t, and cumulative output $X(t) = x$, has a posterior normal distribution of μ, and his expected value can be written as: $E(\mu/m, x, t)$.

The model assumed that the contacts rates of the worker with the employer is given by a Poisson rate λ when he is employed and a Poisson rate δ when he is unemployed. Both rates are exogenous and costless[7]. When a contact is made, the employer screens the worker and depending on the outcome of the screen, the worker decides to take the new job offer or return to his old one. One equilibrium contract for the firm is always to pay the worker his expected marginal product, so that the worker´s pay at each times is given by [Jovanovic (1979a)]:

$$w = E(\mu/m, x, t) \quad (1)$$

While unemployed, the worker gets an unemployment insurance benefit each period (b) as long as he stays unemployed. The present value of becoming unemployed is denoted by U. When a job offer arrives, the worker have an initial wage equal to $E(\mu/m', 0{,}0)$, where m' is the outcome of the screen for the match between the worker and a particular job. These wage offers (from search on the job and search off the job) are drawn from a normal distribution $F(w)$.

$V(w, t)$ is the present value of being on a match. If the cost of changing a job is zero, the worker will change jobs if his on-the-job search yields him an initial wage offer w' such that $V(w', 0) > V(w, t)$. If the reverse inequality holds, he will reject the job offer and continue working on his current job. This defines a reservation wage $\phi(w, t)$, that satisfies the following condition:

$$V(w, t) = V[\phi(w, t), 0] \quad (2)$$

Then if $w' \leq \phi$ the outside offer is rejected.

[7] As the traditional search models this model have search costs, however they do not explain the *mismatch*.

If the current match is a disappointment and acceptable new offers do not arrive, the worker may choose unemployment. Let $\theta(t)$ be the lowest wage the worker will accept on his current job and remain there. At tenure t, a wage lower than $\theta(t)$ will lead the worker to quit and become unemployed. This defines a reservation wage $\theta(t)$ that satisfies the following condition:

$$V[\theta(t), t] = U \quad (3)$$

Finally, ψ is the smallest wage offer that will induce the worker to leave unemployment and accept the job. This defines a reservation wage ψ that satisfies the following condition:

$$U = V[\psi, 0] \quad (4)$$

The present value of becoming unemployed U and the present value of being on a match V are given by:

$$U = \frac{1}{r + \delta[1 - F(\psi)]}\left[b + \delta \int_{\psi}^{\infty} V(x, 0) f(x) dx\right] \quad (5)$$

$$V(w, t) = \max_{\theta, \phi}\left[\int_{t}^{\infty} e^{-r(\tau - t)} \xi(\tau; w, t) d\tau + \int_{t}^{\infty} e^{-r(\tau - t)} \overline{H}(d\tau; w, t, \theta, \psi)\right] \quad (6)$$

Where $\overline{H}(d\tau; w, t, \theta, \psi) = 1 - \int_{\theta(\tau)}^{\infty} H(dw'; \tau, \theta, \psi)$, is the probability that exit takes place on $[t, \tau]$ and $\delta[1 - F(\psi)]$ the hazard rate from unemployment to employment.

As in the search models presented by (Pissarides, 2000), the present value of being unemployed (U) is equal to the discounted income flow of being unemployed (b), plus the gain of receiving an acceptable job offer (equation 5). Moreover, the present value of being on a match (V) is equal to the discounted value of being in the current match or changing for a better job offer plus the discounted value of exiting the current match (equation 6). Using the definitions of reservation wages ($\phi(w, t)$, $\theta(t)$, ψ) and the value of U and V, Jovanovic (1984), solve the system of equations that have a unique solutions for U and V.

Furthermore, Jovanovic (1984, page. 114), defines $p(\tau, w, t)$ as the probability that given wage and tenure (w, t) today, the worker will not have left the current job by time τ, and finds that an increase in the current wage w increases p. The author presents two explanations for this. First, the higher the current wage, the less likely it is that at any future time the wage will fall to the point $\theta(.)$, where $V[\theta(t),t] = U$, at which point the worker will choose unemployment. Moreover, the higher the wage is today, the higher would be in the future, and the less likely it is that the worker will get a better job offer by another firm. This implies that a worker with a higher wage and good match is less likely to leave the current job, while a worker with a low wage or bad match is more likely to leave and search for a better job offer. Second, Jovanovic`s (1984) model predict that a worker would give up a job with a stable wage in favor of a job with a lower initial wage, as long as the new job offered the possibility of wage growth. This prediction has been also found in human capital models (Mincer, 1974). Therefore Jovanovic (1984) model predict that the job to job hazard rate, ($\hat{\lambda}(w,t) \equiv \lambda\left[1 - F\left(\phi(w,t)\right)\right]$, probability of accepting a job offer while being employed), decreases with w and increases in t. This means that if the worker´s wage is held constant while the worker´s productivity grows (because of his tenure), his alternative wage offer will rise and he will be more likely to leave his current job for a better job offer. Therefore, the model implies that workers who find themselves in a *mismatch* situation (low or stable wage), are more likely to leave the current job until they find an adequate match. Therefore, *skill mismatch* might cause labour turnover.

4. Data

To check the relation between *skill mismatch* and labour turnover in Colombia, we build a measure of labour turnover using the Integrated Record of Contributions to Social Security (PILA). PILA is a panel that links information about employers and employees. These administrative records contain information about all hirings and separations[8] made by a formal firm; that is, firms that report the social payments such as health and pension system of their workers (by law the information is reported directly by each firm to the Minister of Health every month). Therefore, we are not covering informal firms.

To focus just on firms, we exclude from our analysis those who are self-employed and are reported in PILA. This account for around 16.2% of the total employment in PILA during 2017. Using PILA we are able to follow around 369.000 firms and 9.420.000 workers in the

[8] We do not have information about voluntary or involuntary separations.

last year[9]. This data allows us to build the worker reallocation rate for the main 23 cities in Colombia in both a monthly and quarterly frequency for the period 2009-2017.

We build the variable Worker Reallocation rate (WR) as our proxy of labour turnover, which refers to the number of workers who change their labour status during a period of time. According to Davis & Haltiwanger (1999), this variable is based on hirings and separations measures as follows (the next notation follows Flórez, Morales, Medina, and Lobo, 2017 and Morales and Medina, 2016):

Assume firm j_t is a set of business-establishments with at least two employees. An individual i_{jt} is an employee observed in the payroll of firm j at period t. Given that we have employer-employee information, we can compute hirings (h_{jt}) as the set of particular employees observed in a given time that were not observed before. Similarly, separations (s_{jt}) are generated as the specific employees found in the previous periods that were not observed in the current one. Then the set of hirings, separations, and stayers (k_{jt}) in a firm j in the period t is defined as:

$$\mathrm{h_{jt}} = \{\mathrm{i}:\ \mathrm{i_t} \in \mathrm{j_t} \ \text{ and } \ \mathrm{i_t} \notin \mathrm{j_{t-1}}\}$$
$$\mathrm{s_{jt}} = \{\mathrm{i}:\ \mathrm{i_t} \notin \mathrm{j_t} \ \text{ and } \ \mathrm{i_t} \in \mathrm{j_{t-1}}\}$$
$$\mathrm{k_{jt}} = \{\mathrm{i}:\ \mathrm{i_t} \in \mathrm{j_t} \ \text{ and } \ \mathrm{i_t} \in \mathrm{j_{t-1}}\}$$

We can calculate aggregate measures by cities, taking summations of all these previous sets. Therefore, the aggregate flows of hirings ($\mathrm{H_{A,t}}$) and separations ($\mathrm{S_{A,t}}$) for city A are represented as:

$$\mathrm{H_{A,t}} = \textstyle\sum_{\mathrm{j} \in \mathrm{A}} \mathrm{h_{jt}};\ \ \mathrm{S_{A,t}} = \textstyle\sum_{\mathrm{j} \in \mathrm{A}} \mathrm{s_{jt}}$$

To simplify the exposition, we omit the subindex A. Thus, worker reallocation is given by:

$$WR_t \equiv H_t + S_t$$

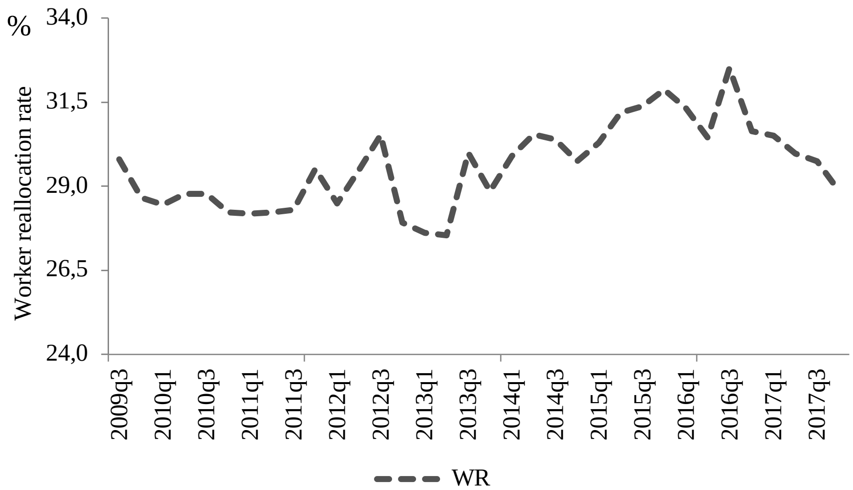
All the measures discussed can be expressed as rates, in which case following Davis et al. (1996a, 1996b), we divide by a measure of employment level of firm j at time t defined

[9] Given the information used in PILA, we are able to follow the change in the number of workers by firm. This measure include workers of both full or part-time jobs, as long as they contribute to the pension or health system.

as: $\mathrm{x_{jt}} = (\mathrm{e_{jt}} + \mathrm{e_{jt-1}})/2$. Therefore, for using this notation, the employment level of any city A can be defined as $\mathrm{X_{A,t}} = \sum_{j \in A} \mathrm{x_{jt}}$. **Figure 2** presents the dynamic across time of the WR rate for the average of 23 main cities in Colombia. As we can observed, the WR rate is around 30% across all the analyzed period. It presents an increasing trend from 2009 to 2015, with a decreasing trend in the last year and a half.

Figure 2: Worker Reallocation Rate (average 23 cities)

(2009-2017 quarterly)



***Source*:** Author's calculations based on PILA.

The second variable we need to build is a measure of *skill mismatch*. Given that PILA does not have information about the level of qualification required by each position in each firm, and we do not have information about the level of education of each worker, we cannot use this data to measure the *skill mismatch*. Therefore, to build our measure of *skill mismatch* by the 23 cities in Colombia, we use the official Colombian Household Survey (GEIH). This is a cross-section survey applied by the Official Statistics Bureau of Colombia (DANE) in a monthly frequency. To make sure that both sources are comparable, we keep only individuals who contribute to the social system payments and work in a firm with two or more employees; that means salaried workers. To have a representative database at the city and occupational level we build our database in a quarterly frequency, both with PILA and GEIH.

In the literature, *skill mismatch* is defined as the difference between the years of education that an individual has and the years of education required by his occupation. There are three ways to measure the educational level required by an occupation (Dolton & Vignoles, 2000; Madrigal, 2003; Mora, 2008; Ramos & Sanromá, 2013). The first one is called "*job analysis*", it consists in determining the educational level required based on the international occupational classifications made by professionals. The second method is called "*self-assessment*" approach, it is based on the assumption that the person currently doing the job can best determine the level of education required[10]. Finally, the "*realized matches*" method, uses information on the educational level of people who exercise a given occupation as the measure of the education requirements, such as the mode, mean etc. (Leuven & Oosterbeek, 2011).

In this paper, we use the *"realized matches"* method. We take individuals with the same occupation and calculate the mode of the years of education. That would be a proxy of the education requirements in each occupation. Then a person is in a *mismatch* situation when his level of education is different from the mode of the educational level of those who exercise his occupation. Moreover, a person is overqualified when his level of education is higher than the mode and underqualified when it is lower (Kiker, Santos, and De Oliveira 1997).

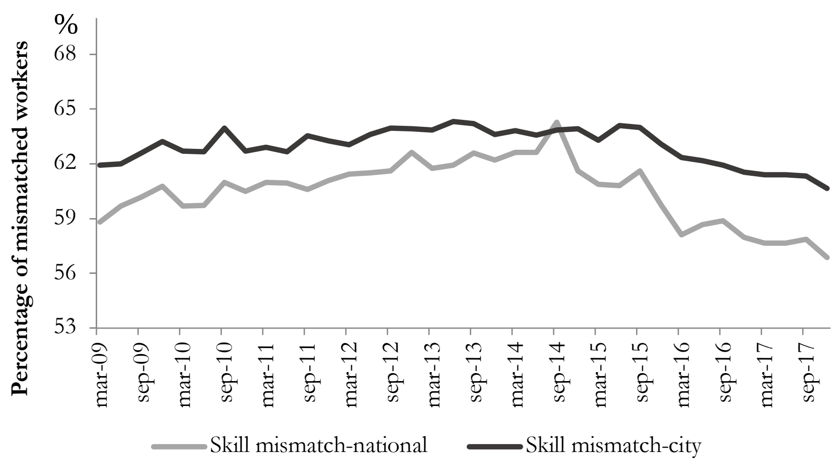
The best measure of *skill mismatch* is to use a measure of the educational level required for all occupations reported by DANE (at the two-digit levels) [11] at each city. These would allow us to capture the heterogeneity in the educational level requirement as the result of the different level of productivity by cities. However, given the limitation of our data, we cannot get a representative measure of the education level requirement at this level of disaggregation.[12] Considering this, we propose two measures of *skill mismatch*. In the first one, we assume that the educational level required by an occupation does not change by cities. Therefore, we use the mode of the years of education at the national level for all

[10] Given the information used in PILA, we are able to follow the change in the number of workers by firm. This measure include workers of both full or part-time jobs, as long as they contribute to the pension or health system.
[11] Another self-reported way to measure *mismatch* is using the calculation of the underemployment by competences used by DANE. However, in this paper, we prefer to avoid the use the self-reported methods.
[12] According to the literature (Verbeek & Nijman, 1992) a that a representative aggregate measure, such as the mode, can be obtained al least with 100 or more observations per group; to eliminate any possible bias in our estimates.

occupations at the two-digit levels (i.e. 82 national modes of all occupations in each period). For each city, we calculate the level of *skill mismatch* as the proportion of people whose level of education do not match the level of education required (at the national level) in his occupation, we call this measure as "*skill mismatch-national*".

Figure 3: *Skill mismatch* comparison between measures

(2009-2017 quarterly)



Source**:** Author's calculations based on the household survey GEIH.

The second measure of *mismatch*, tries to capture the heterogeneity in the educational level required by cities. To do so we have to aggregate some occupations. For this we use the aggregate occupation groups defined by DANE (2015), according with this classification we have 7 occupation groups[13] (i.e. 7 occupation modes for each city- that is 161 modes in each period). Therefore, for each city, we calculate the level of *skill-mismatch* as the proportion of people whose level of education does not match with the level of education required (at the city level) by his occupation, we call this measure as "*skill mismatch-city*".

Figure 3 shows a comparison between the two measures proposed. The figure presents a gap between the two measures of *skill mismatch.* The high *mismatch* is observed when using the *skill mismatch-city*. However, the dynamics across time of both measures is very similar,

[13] According to the classification used by DANE, seven major groups can be defined: 1. Professionals, technicians and assimilated workers; 2. Directors and senior public officials; 3. Administrative staff and assimilated workers; 4. Merchants and sellers; 5. Service workers; 6. Agricultural and forestry workers, fishermen and hunters; and 7. Non-agricultural workers, drivers of machines and vehicles of transport and assimilated workers.

with a reduction in the level of *mismatch* since 2015. Notice that a similar trend across time is observed when comparing the dynamics of the average of WR rate (see Figure 2).

Table 1: Summary statistics average 2009-2017

Variable	Obs	Mean	Std. Dev.	Min	Max
Labour market characteristics					
Worker reallocation rate	828	31.21	6.06	16.75	56.91
Hiring rate	828	16.56	3.60	6.91	38.58
Separation rate	828	14.65	3.21	7.00	26.10
% mismatched workers (*skill mismatch-city*)	828	62.97	4.88	49.65	73.71
% overqualified workers (*skill mismatch-city*)	828	29.94	5.56	15.48	50.16
% underqualified workers (*skill mismatch-city*)	828	33.04	5.58	19.65	51.12
% mismatched workers (*skill mismatch-national*)	828	62.02	4.28	50.57	72.07
% overqualified workers (*skill mismatch-national*)	828	30.11	4.75	17.58	45.68
% underqualified workers (*skill mismatch-national*)	828	31.91	4.57	19.06	45.74
Employees' average potential experience (years)	828	19.39	0.68	17.79	21.40
Employees' average tenure (months)	828	71.32	9.65	46.43	107.22
% workers with fixed-term contract	828	21.49	4.10	10.98	33.10
% workers with higher education	828	22.35	4.49	12.85	38.85
Demographic characteristics					
% male workers	828	57.75	2.48	50.35	65.43
% married workers	828	26.46	5.12	7.61	45.08
% workers under 25 years of age	828	20.03	3.07	10.93	30.86
% workers between 26-40 years of age	828	43.82	2.35	35.70	50.31
% workers between 41-50 years of age	828	20.50	2.06	12.27	28.33
% worker over 51 years of age	828	15.65	1.85	10.77	21.41

***Source*:** Author's calculations based on the household survey GEIH.

Additional to the WR and measures of *skill mismatch* by cities, we build (using the GEIH) other variables that capture the demographic and labour market characteristics of each city such as: sex, marital status, age, workers with higher education, potential experience[14], tenure, whether the person has a fixed-term contract or not, and the economic sector[15]. For the categorical variables, the aggregate measure is the share of workers in each category with

[14] Potential experience is calculated as years of age minus years of education minus 6.

[15] The variable sex takes the value 1 if the worker is a man and 0 otherwise, marital status takes the value 1 if the worker is married and 0 otherwise, the variable age is divided into categories (25 or less, 26-40, 41-50, 51 or more) and takes the value 1 if the worker belongs to a particular category and 0 otherwise, and the variable educational level takes the value 1 if the worker has a higher school degree and 0 otherwise.

respect to the total salaried workers, and for continuous variables it is the mean value for each city.

Table 1 shows the summary statistics for all of our variables of interest, for the period 2009-2017. For the analyzed period the average-hiring rate (16.5%) is greater than the average-separation rate (14.6%), which indicates good performance of the economy. The worker reallocation rate (WR) is around 31.2% with a minimum value of 16.7% and a maximum value of 56.9%. This implies a good variation across cities with a standard deviation of 6.0%. The *skill mismatch* (using the "*skill-mismatch-city*" definition) is on average 62.97%, with 29.94% of overqualified workers and 33.04% underqualified. As we mentioned previously, this measure is high when compared to some developed and developing countries (**see Figure 1**). Moreover, we find that on average the salaried employed workers have a tenure of 71.3 months (5.9 years) where around 21.5% of workers have a fixed-term contract and 22.3% are educated. In terms of the demographic characteristics, we find that 57.7% of the salaried workers are men, 26.6% are married and the majority of the sample is composed of people between 26 and 40 years of age.

5. Empirical strategy and results

5.1 Panel fixed effects approach

In this section, following the prediction of the job matching model presented by Jovanovic (1984), we explore the relationship between the share of *skill-mismatched* workers and the worker reallocation rate across the main 23 cities in Colombia. In order to validate the hypothesis that cities with a higher proportion of *mismatched* workers have higher worker reallocation rates, we estimate panel fixed effects regression by cities. The equation that we estimate can be represented as:

$$WR_{it} = \alpha + Mismatch_{it}\beta + X_{it}'\theta + \gamma_i + \delta_t + \varepsilon_{it} \quad (7)$$

Where WR_{it} is the worker reallocation rate in a city i at period t, $Mismatch_{it}$ is the share of mismatched workers of each city i at period t. Vector X_{it} contains demographic and labour market characteristics of each city such as: the share of workers who are married, the share of workers with a fixed-term contract, the average tenure, the average potential experience of workers, among others. γ_i is a city-specific fixed effect, δ_t stands for time fixed effect, and ε_{it} is the error term. The parameter of interest is β, which captures how the WR change when the *mismatch* measure increases by one unit.

We estimate the model using a within transformation in order to eliminate the constant unobserved heterogeneity which is specific to each city. Given that unobservable common factors are uncorrelated with the explanatory variables, we are confident that our coefficients are not biased. However, the possible presence of cross-sectional dependence between cities may bias our standard errors, therefore, statistical inference based on such standard errors could be invalid (Hoechle, 2007).[16]

Since we are analysing cities in the same country, it is easy to see that these units are likely to be subject to both observable and unobservable common "shocks" that may be explained by country norms, neighbourhood effects, and economic interdependence. To deal with this cross-sectional correlation, we make the correction of standard errors proposed by Driscoll and Kraay (1998), which uses nonparametric methods to estimate a covariance matrix that is consistent to heteroskedasticity, autocorrelation, and is robust to general forms of cross-sectional dependence.

Table 2 presents the results of the estimation of equation (7) using both of our *mismatch* definitions. Columns (1) and (2) shows the results with the *skill mismatch-city* measure and columns (3) and (4) shows the results with *skill mismatch-national* measure. To simplify the interpretation of the coefficients, all variables are standardized; therefore, the effects are interpreted as changes of WR_{it}, expressed in terms of standard deviations, as a result of an increment in one standard deviation of the independent variable. In the second column of each option we include the correction of the standard errors as proposed by Driscoll and Kraay (1998).

[16] Using cluster–robust standard errors will not help here because the correlations across groups of cross sections take nonzero values (De Hoyos & Sarafidis, 2006).

Table 2: Fixed-effects panel estimation by cities

Variables	*Skill mismatch-city* (1) WR Without D-K	(2) WR With D-K	*Skill mismatch-national* (3) WR Without D-K	(4) WR With D-K
Labour market characteristics				
Mismatched workers	0.1276***	0.1276***	0.1081**	0.1081**
	(0.043)	(0.044)	(0.044)	(0.048)
Potential experience	0.2386***	0.2386**	0.2228***	0.2228**
	(0.081)	(0.089)	(0.081)	(0.089)
Tenure	0.1122*	0.1122	0.1149*	0.1149
	(0.058)	(0.080)	(0.058)	(0.080)
Workers with fixed-term contract	0.1121***	0.1121*	0.1124***	0.1124*
	(0.039)	(0.055)	(0.039)	(0.056)
Males	-0.0554	-0.0554	-0.0532	-0.0532
	(0.040)	(0.058)	(0.040)	(0.058)
Married	0.1520**	0.1520	0.1512**	0.1512
	(0.073)	(0.088)	(0.073)	(0.090)
Higher education	0.1342*	0.1342	0.1251*	0.1251
	(0.073)	(0.092)	(0.073)	(0.092)
Demographic characteristics				
Age 25 or less (ref)				
Age 26-40	-0.1047*	-0.1047*	-0.0965*	-0.0965*
	(0.054)	(0.050)	(0.055)	(0.049)
Age 41-50	-0.1984**	-0.1984**	-0.1880**	-0.1880**
	(0.077)	(0.084)	(0.077)	(0.085)
Age 50 or more	-0.3080***	-0.3080***	-0.2933***	-0.2933***
	(0.099)	(0.101)	(0.099)	(0.102)
Constant	1.0157***	1.0157***	1.0117***	1.0117***
	(0.147)	(0.082)	(0.147)	(0.082)
Observations	828	828	828	828
R-squared within	0.3134		0.3109	

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. *Skill mismatch-national*: *mismatch* calculated based on national modes at the two-digit level. The dependent variable is WR in all the specifications. All independent variables represent proportions or means by city. Specifications (2) and (4) are corrected with Driscoll-Kraay standard errors (D-K).

Finally, to guarantee that our *mismatch* measures are representative at the occupational levels explored in this study, we eliminate all occupations with less than 100 individuals by cities in each quarter[17] (Verbeek & Nijman, 1992). This correction eliminates any possible bias in

[17] There are some occupations that are not representative even if we use the sample with all occupied workers instead of salaried workers. This is the case of occupations such as agriculture since the analysis is only for urban zones. Then, from the sample of salaried workers, we lose a 10% of occupations that become unrepresentative.

our *mismatch* estimates[18].Columns (2) and (4) in Table 2 show that, after using the standard errors correction, the proportion of *mismatched* workers has a positive and significant effect on the worker reallocation rate (WR). One standard deviation of increment in the proportion of *mismatched* workers increases the WR rate around 0.12 or 0.11 standard deviations. These results indicate that there is not an important difference between the two measures of *mismatch*. Moreover, cities where workers have a high potential experience and fixed-term contract present a higher WR rate. Then WR rate increase 0.11 standard deviations per standard deviation of increment in the fixed-term contract. Variables such as proportion of males, married and higher education are not significant after the standard errors correction. Finally, the demographic characteristics are important to determine the WR rate. As is suggested by the literature, cities with elderly workers have a lower WR rate than cities with younger workers.

Table A2 from **Appendix** shows the panel estimation when we separate the *mismatched* measure between overqualified and underqualified workers. The results show that the proportion of overqualified workers is significant, while the proportion of underqualified workers is not. Therefore, the *mismatch* effect is explained mainly by the overqualified workers effect (as is also found in the literature, see Mora (2008), Alba-Ramirez (1993), among others). Then, one standard deviation of increment in the proportion of overqualified workers increase the WR rate around 0.21 or 0.18 standard deviations. This results is consistent with the job matching model proposed by Jovanovic (1984), which suggests that if a workers is not satisfied with his current job, he would prefer voluntary to leave his job until he finds a better match.

As we mentioned in the data description, the worker reallocation rate is the sum of separation and hiring rates, therefore, we can separate both measures and estimate equation (7) with each of his components. **Table 3** shows the results for both measures of *mismatch*. The coefficient of mismatched workers is positive and significant in the case of separation rates while it is not when using the hiring rate. This result is in line with the job matching theory,

[18] To make sure our results do not change with the exclusion of the occupations that are not representative in the *skill mismatch* measure, we estimate the same regression as in Table 2 with the full sample. Table A1 from Appendix present these results. In this case, we find that the relation between the proportion of *mismatched* workers and labour turnover is positive and significant.

where the *mismatch* affects the worker reallocation rate only through separations. Notice that these results are very similar for both measures of *mismatch*. Therefore one standard deviation of increment in the proportion of *mismatch* workers increases the separation rate around 0.15 standard deviations.

Table 3: Fixed-effects panel estimation by cities with separations (*S*) and hirings (*H*)

Variables	*Skill mismatch-city*		*Skill mismatch-national*	
	(1) *S* With D-K	(2) *H* With D-K	(3) *S* With D-K	(4) *H* With D-K
Labour market characteristics				
Mismatched workers	0.1533***	0.0780	0.1485**	0.0494
	(0.047)	(0.054)	(0.061)	(0.054)
Potential experience	0.1258	0.2895***	0.1074	0.2792**
	(0.091)	(0.101)	(0.092)	(0.102)
Tenure	0.0286	0.1634	0.0315	0.1654
	(0.052)	(0.127)	(0.053)	(0.127)
Workers with fixed-term contract	0.0999*	0.0996*	0.0997*	0.1003*
	(0.057)	(0.053)	(0.058)	(0.053)
Demographic characteristics				
Age 25 or less (ref)				
Age 26-40	-0.0998	-0.0873	-0.0874	-0.0844
	(0.059)	(0.067)	(0.059)	(0.069)
Age 41-50	-0.1324	-0.2159**	-0.1184	-0.2109**
	(0.089)	(0.090)	(0.091)	(0.092)
Age 50 or more	-0.2057	-0.3349***	-0.1886	-0.3254**
	(0.123)	(0.114)	(0.125)	(0.117)
Constant	1.1276***	0.7039***	1.1226***	0.7015***
	(0.085)	(0.103)	(0.085)	(0.103)
Observations	828	828	828	828

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. *Skill mismatch-national*: *mismatch* calculated based on national modes at the two-digit level. All independent variables represent proportions or means by city. All specifications are corrected with Driscoll-Kraay standard errors (D-K). We also include the proportion of males, married and higher education variables by city however once using the DK correction they are no longer significant.

In order to take into account the productive structure or specific labour demand of each city, we control for the occupation rate by the level of occupations at one-digit code and the occupation rate by the economic sector at each city. This allow us to separate the effect of any specific demand observed in a particular city and the effect of the *skill mismatch* itself.

Table 4 presents the results of the estimation with these additional controls[19]. Column (1) present the results when controlling by occupations, and column (2) when controlling by economic sectors. As we can see, the variable of *skill mismatch* is still significant, and the coefficients are higher than those found in Table 2. However, for this case, the proportion of workers with fixed term contract by city are no longer significant.

Moreover, another variable we control is the institution of the minimum wage. Even though, the minimum wage is determined at national level, its level of restrictiveness (when analysing the relation between the minimum wage ratio and the wage at the 70th percentile) is different across cities. For example, Arango & Flórez (2017) have found a strong heterogeneity in the level of restrictiveness of the minimum wage in Colombia; using these variation the authors find that for cities like Popayán, Villavicencio, Neiva y Florence, the effect of the minimum wage in the informality outcome is higher when comparing with the rest of cities. Column (3-5) control for this minimum wage ratio. Column (3) present the results when using the WR rate. In this case, as we expect, cities with a restrictive minimum wage (high minimum wage ratio) present low WR rate (similar results are found by Flórez, Morales, Medina, & Lobo, 2017). The channel where we can see this effect is through a lower hiring rate- H (see column (5)). Then, when controlling by the minimum wage institution we find that the effect of *skill mismatch* is still positive and significant, especially when studying the WR rate and the separation rate (column (3) and (4)). Similar results were found when using measure *of skill mismatch-national*, see Table A3 in appendix.

5.2. Robustness check[20]

One of the robustness check we can implement in our empirical exercise is to change the criteria we use to define when a person is in a *mismatch* situation. Remember, so far we define a *mismatch* if the years of education of the individual are different from the mode of

[19] We also control by the size of the firm, including the proportion of workers who work in firms with 2 to 5 employees, 6 to 100 employees, and more than 100 employees. Our results show that cities with a high proportion of workers in large firms have a higher worker reallocation rate. The results are upon request.

[20] We also did the exercise of building the WR using data from GEIH. We have the change in employment and we can calculate hirings as individuals who have a tenure less than 3 months. Then, separations can be calculated as the difference between the change in employment and hirings. The mismatch variable is positive and significant in the case of the skill mismatch-national measure, but with the skill mismatch-city measure, the coefficients are not significant. The WR measure built with data from GEIH is less precise than WR from PILA. The first one is estimated using a sample while the second one utilizes the real WR reallocation built from administrative records. For this reason, we prefer to use the WR from PILA which is more accurate.

the years of education of his occupation. However, we can define a *mismatch* if the years of education of the individual are different from the mode plus 0.5 standard deviations of his occupation. In this case, the probability of being in *mismatch* is lower than in the previous case.

Table 4: Fixed-effects panel estimation by cities with additional controls

Variables	*Skill mismatch-city*				
	(1) WR	**(2)** WR	**(3)** WR	**(4)** S	**(5)** H
	With D-K	**With D-K**	**With D-K**	**With D-K**	**With D-K**
Labour market characteristics					
Mismatched workers	0.1611***	0.1514***	0.1326**	0.1700***	0.0715
	(0.032)	(0.040)	(0.050)	(0.054)	(0.060)
Potential experience	0.3506**	0.3091**	0.1830**	0.0831	0.2339**
	(0.129)	(0.113)	(0.083)	(0.081)	(0.098)
Tenure	0.0846	0.0934	0.1589*	0.0265	0.2438*
	(0.073)	(0.075)	(0.081)	(0.050)	(0.1245)
Workers with fixed-term contract	0.0830	0.0784	0.0941	0.0908	0.0773
	(0.052)	(0.061)	(0.068)	(0.063)	(0.065)
Demographic characteristics					
Age 25 or less (ref)					
Age 26-40	-0.1647**	-0.1438**	-0.1093**	-0.0569	-0.1332**
	(0.065)	(0.062)	(0.051)	(0.053)	(0.062)
Age 41-50	-0.2863**	-0.2450**	-0.1264*	-0.0426	-0.1748*
	(0.111)	(0.103)	(0.072)	(0.063)	(0.088)
Age 50 or more	-0.4220***	-0.3885***	-0.2679**	-0.1285	-0.3363**
	(0.141)	(0.126)	(0.103)	(0.105)	(0.123)
Constant	1.1975***	0.9857***	0.4754***	0.2723***	0.5573***
	(0.090)	(0.094)	(0.049)	(0.065)	(0.057)
Institutions					
Δ Minimum wage relative to the 70th percentile			-0.0541**	-0.0374	-0.0577**
			(0.024)	(0.023)	(0.023)
Control by occupation groups	✓				
Control by economic sectors		✓			
Observations	828	828	736	736	736

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. All variables represent proportions or means by area. All specifications are corrected with Driscoll-Kraay standard errors (D-K). We also include three variables by city, the proportion of males, proportion of married workers and proportion of workers with higher education. However after using the DK correction they were no longer significant.

Table 5: Fixed-effects panel estimation by cities changing the definition of mismatch

Variables	*Skill mismatch-city* (1) WR Without D-K	*Skill mismatch-city* (2) WR With D-K	*Skill mismatch-national* (3) WR Without D-K	*Skill mismatch-national* (4) WR With D-K
Labour market characteristics				
Mismatched workers	0.0741**	0.0741**	0.0748*	0.0748
	(0.032)	(0.031)	(0.039)	(0.044)
Potential experience	0.2312***	0.2312**	0.2111***	0.2111**
	(0.081)	(0.092)	(0.081)	(0.091)
Tenure	0.1002*	0.1002	0.1081*	0.1081
	(0.059)	(0.081)	(0.059)	(0.081)
Workers with fixed-term contract	0.1189***	0.1189*	0.1165***	0.1165*
	(0.039)	(0.057)	(0.039)	(0.057)
Males	-0.0593	-0.0593	-0.0492	-0.0492
	(0.040)	(0.056)	(0.040)	(0.059)
Married	0.1337*	0.1337	0.1447**	0.1447
	(0.073)	(0.086)	(0.073)	(0.090)
Higher education	0.1270*	0.1270	0.1169	0.1169
	(0.073)	(0.094)	(0.072)	(0.093)
Demographic characteristics				
Age 25 or less (ref)				
Age 26-40	-0.1097**	-0.1097**	-0.0971*	-0.0971**
	(0.054)	(0.049)	(0.055)	(0.046)
Age 41-50	-0.1996***	-0.1996**	-0.1848**	-0.1848**
	(0.077)	(0.084)	(0.077)	(0.085)
Age 50 or more	-0.3001***	-0.3001***	-0.2804***	-0.2804**
	(0.099)	(0.104)	(0.099)	(0.104)
Constant	1.0116***	1.0116***	0.9899***	0.9899***
	(0.147)	(0.084)	(0.148)	(0.082)
Observations	828	828	828	828
R-squared Within	0.3101		0.3087	

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. *Skill mismatch-national*: *mismatch* calculated based on national modes at the two-digit level. The dependent variable is WR in all the specifications. All independent variables represent proportions or means by city. Specifications (2) and (4) are corrected with Driscoll-Kraay standard errors (D-K).

Table 5 presents the results changing the definition of *mismatch*. The *mismatch* coefficient in the regression with the measure *skill mismatch-city* is still positive and significant. However, comparing the results with those found in Table 2, the coefficients are lower, therefore, one standard deviation of increment in the proportion of *mismatched* workers increases the WR rate around 0.07 standard deviations. The rest of variables still significant and with the expected signs. However, when we analyse the results using the *skill mismatch-national* measure the coefficient with the Driscoll-Kraay correction is no longer significant. These results may be explained due to the fact that this measure presents a lower variation across cities, then, changing the strictness of the definition may reduce more his variation (see Table 1).

Finally, we explore an additional approach, also suggested in the literature, to capture the relationship between *skill mismatch* and labour mobility. In this case, we use individual data instead of a panel data by cities. In this case, we are not able to control for the unobserved heterogeneity. However, using the pool of individual data, we are able to use another proxy of labour turnover, widely used in the literature (Hersch, 1991, 1995; McGuinness & Wooden, 2007; Robst, 1995; Verhaest & Omey, 2006). This variable is the "quit intention" and is built using the following question from the GEIH: Do you want to change your job? Therefore, the dependent variable takes the value 1 if the worker wants to change his job and 0 otherwise. The independent variables are the same as in the previous exercises but they are defined for each individual instead of cities.

Table 6 presents the marginal effects of a probit estimation using the *skill mismatch-city* and *skill mismatch-national* measures. Moreover, given that we are using only the information from the GEIH we can separate the individuals into salaried workers and non-salaried workers. Again, in this case, the *mismatch* variable is positive and significant but only for those salaried workers. Then a salaried worker in a *mismatch* situation is more likely to quit his job. The rest of the variables are also significant and, in general, lead to the same conclusions as when using the fixed effect model approach. The result for those non-salaried workers is expected to be not significant, given the nature of non-salaried workers who are usually self-employed and informal workers.

Table 6: Pool regression: marginal effects from a probit

Variables	*Skill mismatch-city* (1) Quit intention Salaried	(2) Quit intention Non-salaried	*Skill mismatch-national* (3) Quit intention Salaried	(4) Quit intention Non-salaried
Labour market characteristics				
Mismatched workers	0.0095***	-0.0013	0.0119***	0.0015
	(0.001)	(0.001)	(0.001)	(0.001)
Potential experience	0.0014***	-0.0023***	0.0013***	-0.0022***
	(0.000)	(0.000)	(0.000)	(0.000)
Tenure	-0.0009***	-0.0008***	-0.0009***	-0.0008***
	(0.000)	(0.000)	(0.000)	(0.000)
Workers with fixed-term contract	-0.0012	-0.0723***	0.0004	-0.0721***
	(0.001)	(0.007)	(0.001)	(0.006)
Males	-0.0147***	0.0232***	-0.0137***	0.0246***
	(0.001)	(0.001)	(0.001)	(0.001)
Married	-0.0385***	-0.0615***	-0.0389***	-0.0637***
	(0.001)	(0.001)	(0.001)	(0.001)
Demographic characteristics				
Age 25 or less (ref)				
Age 26-40	-0.0243***	0.0418***	-0.0241***	0.0427***
	(0.002)	(0.002)	(0.002)	(0.002)
Age 41-50	-0.0449***	0.0504***	-0.0451***	0.0505***
	(0.003)	(0.004)	(0.003)	(0.004)
Age 50 or more	-0.0954***	-0.0046	-0.0951***	-0.0056
	(0.004)	(0.005)	(0.004)	(0.005)
Higher education	-0.0371***	-0.0964***	-0.0390***	-0.0981***
	(0.001)	(0.003)	(0.001)	(0.003)
Observations	1,497,733	1,282,602	1,512,650	1,363,030
Pseudo R^2	0.0488	0.0532	0.0498	0.0532

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. We use robust standard errors correction.

6. Conclusion

We show that there exists a relationship between *skill mismatch* and labour reallocation in a developing country such as Colombia, as suggested by Jovanovic`s (1979a, 1979b, 1984) job matching theory. Using an aggregate panel data at the city level (23 main cities in Colombia), with a quarterly frequency for the period 2009-2017, we find that one standard deviation of increment in the proportion of *mismatch* workers increases the WR rate around 0.12 standard deviations. Moreover, we find that the WR rate increase 0.11 standard deviations per standard deviation of increment in the fixed-term contract. Variables such as proportion of workers who are males, married and with higher education are not significant after the Driscoll-Kraay standard errors correction. Finally, the demographic characteristics are important to determine the WR rate. As is suggested by the literature, cities with elderly workers have a lower WR rate than cities with younger workers.

We also find that the positive effect of *skill mismatch* and labour reallocation is explained mainly by the increase in separations, in line with our theoretical framework.. Therefore, one standard deviation of increment in the proportion of mismatch workers increases the separation rate around 0.15 standard deviations. The effect on hirings is not significant. We also control for the productive structure of each city, using the occupation rate by the level of occupations at one-digit code and the occupation rate by the economic sector at each city. Even controlling for the different productive structure of each city, we find that the *skill mismatch* is still significant and positive.

Furthermore, we control for the level of restrictiveness of the minimum wage in each city (measured as the relation between the minimum wage ratio and the wage at the 70th percentile). We find that cities with a restrictive minimum wage (high minimum wage ratio) present low WR rate (similar results are found by Flórez, Morales, Medina, & Lobo, 2017). Then, when controlling for the minimum wage institution we find that the effect of *skill mismatch* is still positive and significant, especially when studying the WR rate and the separation rate.

Finally, we capture the relationship between *skill mismatch* and labour mobility using individual data, instead of panel data. In this case, we are not able to control for the unobserved heterogeneity, but we can use a widely used variable in the literature, the "quit intention". In this case, the *mismatch* variable is positive and significant for salaried workers. Then a salaried workers in a *mismatch* situation is more likely to quit his job. The rest of the variables are also significant and, in general, lead to the same conclusions as when using the fixed effect model approach.

As we mentioned before, the *skill mismatch* is costly for both firms and workers. We find that it increases the labour turnover in the labour market, causing a negative effect on the human capital accumulation in a particular firm, with a potential negative effect on the level of productivity of the firms. Therefore, we need to develop strategies that improve the adequate qualification of workers, as well as the search process for both workers and firms in order to reduce the level of *mismatch* in a developing country like Colombia.

References

Alba-Ramirez, A. (1993). Mismatch in the Spanish labor market: overeducation? *Journal of Human Resources*, *28*(2), 259–278.

Alba-Ramírez, A., & Blázquez, M. (2003). Types of job match, overeducation and labour mobility in Spain. *Overeducation in Europe*, 65–92.

Allen, J., & Van der Velden, R. (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers, 55*(3), 434–452.

Arango, L. E., & Flórez, L. A. (2017). Informalidad laboral y elementos para un salario mínimo diferencial por regiones en Colombia. *Borradores de Economía*, (1023).

Badillo-Amador, L., & Vila, L. E. (2013). Education and skill mismatches: wage and job satisfaction consequences. *International Journal of Manpower*, *34*(5), 416–428.

Becker, G. (1964). Human Capital. : *: National Bureau of Economic Research (NBER)*.

Büchel, F., & Mertens, A. (2004). Overeducation, undereducation, and the theory of career mobility. *Applied Economics*, *36*(8), 803–816.

Castillo, M. (2007). Desajuste educativo por regiones en Colombia: ¿Competencia por salarios o por puestos de trabajo? *Cuadernos de Economía*, *26*(46), 107–145.

DANE. (2015). Metodología General Gran Encuesta Integrada de Hogares. Retrieved from https://www.dane.gov.co/files/investigaciones/fichas/empleo/metodologia_GEIH-01_V9.pdf

Davis, S. J., & Haltiwanger, J. (1999). Handbook of Labor Economics. In *Handbook of Labor Economics* (Vol. 3, pp. 2711–2805). Elsevier.

Davis, S. J., Haltiwanger, J. C., & Schuh, S. (1996). *Job creation and destruction*. MIT Press.

De Hoyos, R., & Sarafidis, V. (2006). Testing for Cross-Sectional Dependence in Panel-Data Models. *The Stata Journal: Promoting Communications on Statistics and Stata, 6*(4), 482–496.

Di Pietro, G., & Urwin, P. (2006). Education and skills mismatch in the Italian graduate labour market. *Applied Economics*, *38*(1), 79–93.

Dolton, P., & Vignoles, A. (2000a). The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review*, *19*(2), 179–198.

Dolton, P., & Vignoles, A. (2000b). The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review*, *19,* 179–198.

Domínguez, J. (2009). Sobreeducación en el mercado laboral urbano de Colombia para el año 2006. *Sociedad y Economía*, (16), 141–160.

Driscoll, J., & Kraay, A. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics, 80*(4), 549–560.

Flórez, L. A., Morales, L. F., Medina, D., & Lobo, J. (2017). Labour flows across firm´s size, economic sectors and wages: evidence from employer-employee linked panel.

Borradores de Economía, (1013), 1–50.
Frei, C., & Sousa-Poza, A. (2012). Overqualification: permanent or transitory? *Applied Economics*, *44*(14), 1837–1847.
Groot, W., & Maassen van den Brink, H. (2003). The dynamics of skill mismatches in the Dutch labour market. *Overeducation in Europe*, 49–64.
Hartog, J. (2000). *Over-education and earnings: where are we, where should we go? Economics of Education Review* (Vol. 19).
Herrera-Idárraga, P., López-Bazo, E., & Motellón, E. (2013). *Informality and overeducation in the labor market of a developing country*. Barcelona.
Hersch, J. (1991). Education match and job match. *The Review of Economics and Statistics*, *73*(1), 140–144.
Hersch, J. (1995). Optimal "mismatch" and promotions. *Economic Inquiry*, *33*(4), 611–624.
Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *The Stata Journal: Promoting Communications on Statistics and Stata*, *7*(3), 281–312.
Jovanovic, B. (1979a). Firm-specific Capital and Turnover. *Journal of Political Economy*, *87*(6), 1246–1260.
Jovanovic, B. (1979b). Job Matching and the Theory of Turnover. *Journal of Political Economy*, *87*(5, Part 1), 972–990.
Jovanovic, B. (1984). Matching, Turnover, and Unemployment. *Journal of Political Economy*, *92*(1), 108–122. https://doi.org/10.1086/261210
Kampelmann, S., & Rycx, F. (2012). The impact of educational mismatch on firm productivity: Evidence from linked panel data. *Economics of Education Review*, *31*(6), 918–931.
Kiker, B., Santos, M., & De Oliveira, M. (1997). Overeducation and Undereducation: Evidence for Portugal. *Economics of Education Review*, *16*(2), 111–125.
Leuven, E., & Oosterbeek, H. (2011). Overeducation and Mismatch in the Labor Market. In *Handbook of the Economics of Education, volume 4* (p. 283–326. Chapter 3).
Lucifora, C., & Origo, F. (2002). *The economic cost of the skill gap in Europe*.
Madrigal, M. (2003). Una revisión de los métodos de medición del desajuste educativo: ventajas e inconvenientes. *In Hacienda pública y convergencia europea: X Encuentro de Economía Pública, Santa Cruz de Tenerife.* Universidad de la Laguna.
Manacorda, M., & Petrongolo, B. (1999). *Skill Mismatch and Unemployment in OECD Countries*. *Economica* (Vol. 66).
McCall, B. (1990). Occupational Matching: A Test of Sorts. *Journal of Political Economy*, *98*(1), 45–69.
McGowan, M. A., & Andrews, D. (2017). Labor Market Mismatch and Labor Productivity: Evidence from PIAAC Data. In K. T. Solomon W. Polachek , Konstantinos Pouliakas , Giovanni Russo (Ed.), *Skill mismatch in labour markets* (pp. 199–241). Emerald Publishing Limited.
Mcguinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, *20*(3), 387–418.
McGuinness, S., & Wooden, M. (2007). Overskilling, Job Insecurity, and Career Mobility. *Industrial Relations: A Journal of Economy and Society*, *48*(2), 265–286.
Miller, R. (1984). Job Matching and Occupational Choice. *Journal of Political Economy*, *92*(6), 1086–1120.
Mincer, J. (1974). Schooling, experience and earnings. *New York: NBER*.

Mora, J. (2008). Sobre-educación en el mercado laboral colombiano. *Revista de Economía Institucional*, *10*(19), 293–309.
Morales, L., & Medina, D. (2016). Labor Fluidity and Performance of Labor Outcomes in Colombia: Evidence from Employer-Employee Linked Panel. *Borradores de Economía*, (926).
OECD. (2017). *Getting Skills Right: Skills for Jobs Indicators, Getting Skills Right*. OECD Publishing, Paris.
Olitsky, N. (2008). The Procyclicality of Mismatches. *University of Massachusetts-Dartmouth, Mimeo.*
Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT Press.
Quejada, R., & Ávila, J. (2017). Sobreeducación en Colombia: un análisis de los determinantes y desajustes del mercado laboral en un contexto nacional y regional. *Trabajo y Sociedad* , (28), 219–236.
Quintini, G. (2011). *Over-Qualified or Under-Skilled: A Review of Existing Literature*. *OECD Social, Employment and Migration Working Papers*.
Ramos, R., & Sanromá, E. (2013). Overeducation and Local Labour Markets in Spain. *Journal of Economic and Social Geography*, *104*(3), 278–291.
Robst, J. (1995). Career mobility, job match, and overeducation. *Eastern Economic Journal*, *21*(4), 539–550.
Rosen, S. (1972). Learning and experience in the labor market. *Journal of Human Resources*, 236–342.
Rubb, S. (2013). Overeducation, undereducation and asymmetric information in occupational mobility. *Applied Economics*, *45*(6), 741–751.
Rumberger, R. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources*, 24–50.
Serrano, C., & Malo, M. (1996). Desajuste educativo y movilidad laboral en España. *Revista de Economía Aplicada*, *4*(11), 105–131.
Sicherman, N. (1991). Overeducation in the Labor Market. *Journal of Labor Economics, 9*(2), 101–122.
Sicherman, N., & Galor, O. (1990). A Theory of Career Mobility. *Journal of Political Economy*, *98*(1), 169–192.
Sloane, P. J., Battu, H., & Seaman, P. T. (1999). Overeducation, undereducation and the British labour market. *Applied Economics*, *31*(11), 1437–1453.
van der Velden, R., & Verhaest, D. (2017). Are Skill Deficits always Bad? Toward a Learning Perspective on Skill Mismatches. In S. W. Polachek, K. Pouliakas, G. Russo, & T. Konstantinos (Eds.), *Skill Mismatch in Labor Markets* (pp. 305–343). Emerald Publishing Limited.
Verbeek, M., & Nijman, T. (1992). Can Cohort Data be Treated as Genuine Panel Data? In *Panel Data Analysis* (pp. 9–23). Heidelberg: Physica-Verlag HD.
Verdugo, R. R., & Verdugo, N. T. (1989). The Impact of Surplus Schooling on Earnings: Some Additional Findings. *The Journal of Human Resources*, *24*(4), 629.
Verhaest, D., & Omey, E. (2006). The Impact of Overeducation and its Measurement. *Social Indicators Research*, *77*(3), 419–448.
Wen, L., & Maani, S. A. (2019). Job mismatches and career mobility. *Applied Economics, 51*(10), 1010–1024.

Appendix

Table A1: Fixed-effects panel estimation by cities with the full sample

Variables	*Skill mismatch-city* (1) Without D-K	*Skill mismatch-city* (2) With D-K	*Skill mismatch-national* (3) Without D-K	*Skill mismatch-national* (4) With D-K
Labour market characteristics				
Mismatched workers	0.1314***	0.1314***	0.1179***	0.1179**
	(0.043)	(0.045)	(0.044)	(0.048)
Potential experience	0.2395***	0.2395**	0.2231***	0.2231**
	(0.081)	(0.090)	(0.081)	(0.089)
Tenure	0.1125*	0.1125	0.1146*	0.1146
	(0.058)	(0.080)	(0.058)	(0.079)
Workers with fixed-term contract	0.1126***	0.1126*	0.1121***	0.1121*
	(0.039)	(0.055)	(0.039)	(0.056)
Males	-0.0568	-0.0568	-0.0534	-0.0534
	(0.040)	(0.058)	(0.040)	(0.059)
Married	0.1517**	0.1517	0.1522**	0.1522
	(0.073)	(0.088)	(0.073)	(0.090)
Higher education	0.1374*	0.1374	0.1270*	0.1270
	(0.073)	(0.092)	(0.072)	(0.092)
Demographic characteristics				
Age 25 or less (ref)				
Age 26-40	-0.1047*	-0.1047*	-0.0953*	-0.0953*
	(0.054)	(0.050)	(0.055)	(0.048)
Age 41-50	-0.1996***	-0.1996**	-0.1875**	-0.1875**
	(0.077)	(0.085)	(0.077)	(0.085)
Age 50 or more	-0.3099***	-0.3099***	-0.2933***	-0.2933***
	(0.099)	(0.102)	(0.099)	(0.102)
Constant	1.0144***	1.0144***	1.0086***	1.0086***
	(0.147)	(0.082)	(0.147)	(0.082)
Observations	828	828	828	828
R-squared Within	0.3138		0.3117	

Notes: Standard error in parentheses. * p<0.1 **p<0.05 ***p<0.01 *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. *Skill mismatch-national*: *mismatch* calculated based on national modes at the two-digit level. The dependent variable is WR in all the specifications. All variables represent proportions or means by area. Specifications (2) and (4) are corrected with Driscoll-Kraay standard errors (D-K).

Table A2: Fixed-effects panel estimation by cities dividing *mismatch* by overqualified and underqualified workers

Variables	*Skill mismatch-city*		*Skill mismatch-national*	
	(1) **Without D-K**	**(2)** **With D-K**	**(3)** **Without D-K**	**(4)** **With D-K**
<u>Labour market characteristics</u>				
Overqualified workers	0.2183***	0.2183***	0.1807***	0.1807***
	(0.052)	(0.056)	(0.054)	(0.049)
Underqualified workers	0.0433	0.0433	-0.0076	-0.0076
	(0.055)	(0.059)	(0.067)	(0.106)
Potential experience	0.3357***	0.3357***	0.3411***	0.3411***
	(0.084)	(0.097)	(0.092)	(0.119)
Tenure	0.0852	0.0852	0.0873	0.0873
	(0.058)	(0.081)	(0.059)	(0.084)
Workers with fixed-term contract	0.1039***	0.1039*	0.1025***	0.1025*
	(0.039)	(0.057)	(0.039)	(0.056)
Males	-0.0565	-0.0565	-0.0579	-0.0579
	(0.039)	(0.055)	(0.040)	(0.057)
Married	0.1274*	0.1274	0.1505**	0.1505
	(0.072)	(0.087)	(0.073)	(0.090)
Higher education	0.1519**	0.1519*	0.1047	0.1047
	(0.072)	(0.086)	(0.073)	(0.082)
<u>Demographic characteristics</u>				
Age 25 or less (ref)				
Age 26-40	-0.1528***	-0.1528***	-0.1420**	-0.1420**
	(0.055)	(0.050)	(0.057)	(0.051)
Age 41-50	-0.2692***	-0.2692***	-0.2658***	-0.2658**
	(0.078)	(0.089)	(0.082)	(0.101)
Age 50 or more	-0.3966***	-0.3966***	-0.4013***	-0.4013***
	(0.101)	(0.107)	(0.107)	(0.121)
Constant	1.0544***	1.0544***	0.9121***	0.9121***
	(0.146)	(0.082)	(0.151)	(0.097)
Observations	828	828	828	828
R-squared Within	0.3261		0.3168	

Notes: Standard error in parentheses. * p<0.1 **p<0.05 ***p<0.01 *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. *Skill mismatch-national*: *mismatch* calculated based on national modes at the two-digit level. The dependent variable is WR in all the specifications. All variables represent proportions or means by area. Specifications (2) and (4) are corrected with Driscoll-Kraay standard errors (D-K).

Table A3: Fixed-effects panel estimation by cities with additional controls

Variables	*Skill mismatch-national*				
	(1) WR	**(2) WR**	**(3) WR**	**(4) S**	**(5) H**
	With D-K	**With D-K**	**With D-K**	**With D-K**	**With D-K**
<u>Labour market characteristics</u>					
Mismatched workers	0.1478***	0.1320***	0.1013*	0.1709**	0.0181
	(0.041)	(0.042)	(0.056)	(0.066)	(0.062)
Potential experience	0.3530**	0.2922**	0.1683*	0.0651	0.2252**
	(0.133)	(0.114)	(0.085)	(0.082)	(0.100)
Tenure	0.0869	0.0956	0.1617*	0.0292	0.2462*
	(0.073)	(0.074)	(0.081)	(0.051)	(0.123)
Workers with fixed-term contract	0.0841	0.0780	0.0946	0.0900	0.0790
	(0.054)	(0.062)	(0.069)	(0.063)	(0.066)
<u>Demographic characteristics</u>					
Age 25 or less (ref)					
Age 26-40	-0.1546**	-0.1325**	-0.1016*	-0.0408	-0.1347*
	(0.062)	(0.061)	(0.050)	(0.052)	(0.065)
Age 41-50	-0.2844**	-0.2322**	-0.1163	-0.0249	-0.1736*
	(0.108)	(0.104)	(0.075)	(0.063)	(0.093)
Age 50 or more	-0.4206***	-0.3706***	-0.2530**	-0.1095	-0.3282**
	(0.142)	(0.1275)	(0.105)	(0.105)	(0.129)
<u>Institutions</u>					
Δ Minimum wage relative to the 70th percentile			-0.0552**	-0.0380	-0.0591**
			(0.024)	(0.024)	(0.023)
Constant	1.1747***	0.9802***	0.4718***	0.2875***	0.5376***
	(0.132)	(0.093)	(0.051)	(0.072)	(0.055)
Control by occupational groups	✓				
Control by for economic sector		✓			
Observations	828	828	736	736	736

Notes: Standard error in parentheses. * p<0.1, **p<0.05, ***p<0.01. *Skill mismatch-city*: *mismatch* calculated based on modes by city at the one-digit level. All variables represent proportions or means by area. All specifications are corrected with Driscoll-Kraay standard errors (D-K). We also include three variables by city, the proportion of males, proportion of married workers and proportion of workers with higher education. However, after using the DK correction they were no longer significant.


BANCO DE LA REPÚBLICA
COLOMBIA