Wage gaps across Colombian regions: the role of education and informality

Paula Herrera-Idárraga*
AQR Research Group – University of Barcelona and Pontificia Universidad Javeriana.
Enrique López-Bazo
AQR Research Group – University of Barcelona
Elisabet Motellón
Universitat Oberta de Catalunya and AQR Research Group – University of Barcelona

Preliminary version (February 2014) – Please do not quote

Abstract

This paper uses Colombian micro-data to analyze the role of education and informality on regional wage differentials. The hypothesis of this study is that apart from the difference in the endowments of human capital across regions, regional heterogeneity in the incidence of informality may be another important source of regional wage inequality. The results for Colombian regions confirm marked differences in wage distributions between regions and that they differ in the endowment of human capital and more importantly in the incidence of informality. Regional heterogeneity in returns to education is especially intense in upper part of the wage distribution. While heterogeneity in the informal pay penalty throughout the territory is more relevant in the lower part of the wage distribution.

Acknowledgements: The authors gratefully acknowledge financial support from the Spanish Ministry of Science and Innovation, *Plan Nacional de I+D+i*, Project ECO2011-30260-C03-03, and the European Commission under the 7th Framework Programme project SEARCH, SSH-2010-2.2-266834. Paula Herrera-Idárraga also acknowledges financial support from the Agència de Gestiò d'Ajuts Universitaris i de Recerca (AGAUR) (grant FI-DGR 2012).

***Corresponding author:** Paula Herrera-Idárraga. Mailing address: Diagonal, 690, 08034 Barcelona, Spain; telephone number: +0034-934031907; e-mail: paulaherrera@ub.edu.

1. Introduction

Over the past decade, several studies have registered the decline in income inequality for Latin America countries (López-Calva & Lustig, 2009 and Gasparini, Cruces and Tornarolli, 2011). Regional studies are of great relevance, because even in the presence of declining income inequality at national level, important inter-regional disparities may exist. This is so, because socio economic indicators at the national level can often hide significant variances between territories of the same country. This study considers the case of Colombia, a country that despite a decrease in income inequality in the past decade presents one of the highest Gini coefficients of Latin America countries and faces large geographical differences. Colombia shows important disparities in economic and social development among its regions. This implies that an important part of inequality between Colombian individuals may be the consequence of inequality between regions of the country (Bonet and Meisel, 2008; Jourmard and Londono, 2013). In particular, differences in wages deserve attention from a regional perspective as, for example, in 2010 the average gross hourly wage of a small city, such as Cucuta, is only 66% of that paid in Bogotá.

To explain large spatial wage disparities, several explanations have been proposed. One of them emphasizes that wage differences across areas are caused by differences in amenities. For instance, certain areas may have a favorable climate and more access to natural resources. Under this context, wage differentials may be seen as compensated differentials, meaning that some areas may have higher wages to attract workers so to compensate for the lack of amenities (Greenwood et al. 1991). Another explanation is related to the point that differences in wages across regions could reflect spatial differences in the skill composition of the workforce (Combes, Duranton and Gobillon, 2008). Workers with better labor market characteristics tend to sort themselves in areas that concentrate industries with high skill requirements where wages tend to be higher. Associated to this last explanation, the third one is based on agglomeration economies. A larger pool of high skill workers in an area may provide a source of important knowledge spillovers that can lead to productivity gains (Glaeser et al., 1992). Also, labor pooling improves the matching between firms and workers, which could also increase economic efficiency and leads to higher wages (Andersson, Burgess and Lane, 2007).

A number of studies have been devoted at measuring the degree of regional wage gaps and identifying their origin. Blackaby and Murphy (1995) and Duranton and Monastiriotis (2002) analyze the case of Britain, García and Molina (2002), Motellón, López-Bazo and El-Attar (2011), López-Bazo and Motellón (2011) that of Spain and Pereira and Galego (2013) the one of Portugal. These studies center their analysis on the estimation of human capital wage equations and on decomposition analysis. The decomposition analysis is based on the idea that regional wage differentials are the result of the difference in which, characteristics that determines wages are distributed across regions (the characteristics component) and by how different these characteristics are rewarded across space (the coefficients or wage structure component). The extent to which these two components explain regional wage differentials has been of great interest in past studies and their importance in explaining regional wage gaps differ considerably across and within countries. Some studies conclude that the regional wage differentials are mostly due to differences in individual characteristics between regions (Blackaby and Murphy, 1995). Other studies found that a significant part of wage differentials are explained by difference in returns, (Motellón, López-Bazo and El-Attar, 2011 and Pereira and Galego, 2013). While some studies point that both components play an important role (García and Molina, 2002).

To the best of our knowledge almost all studies that analyze regional wage differentials for Colombia are aggregate approaches. These approaches are centered on a single aggregate variable, usually per capita income, at the regional level. For example, Bonet and Meisel (2007) study the convergence in regional income in Colombia, for the administrative division of departments and analyzed the period comprehend between the years 1975-2000, concluding that there is a process of polarization. Bogotá, whose per capita income is more than twice the national mean during the period of analysis, is the regional unit that leads this process. Unfortunately, aggregate approaches hardly say anything about what factors explain regional inequalities. Our study pretends to be one of the few disaggregated approaches for Colombia, which with the use of micro data will try to give some light into what factors accounts for regional income inequalities.

Given the importance of labor market inequality dynamics in explaining the trend in inequality, and since earnings obtained in the labor market are the main sources of income; this paper will be focus on analyzing wage inequality at the regional level. As in previews studies for other countries, special attention is paid to spatial imbalances in the endowment of human capital, and to what extend these differences and the regional heterogeneity in the return to this type of capital may help to explain regional wage gaps. However as a novel and main contribution, this paper will not only focus on the regional differences in the endowments of human capital, but will go further in exploring one important feature of almost all developing countries: the stylized fact that a large proportion of the employed population in Colombia has an informal job. More importantly, recent studies for Colombia have emphasized that informal jobs are not equally distributed across the main metropolitan areas of the country (Galvis, 2012). In Colombia some cities have informality rates of around 70% while others have rates of about 50%. In addition, we build on the results in the study by Ortiz, Uribe and Badillo (2008), which indicates that the Colombian labor market is segmented in two dimensions. An intra-regional or scale segmentation, which is mainly due to the restrictions on the access to physical and human capital that limited the possibility of expansion of firms to a larger scale. This type of segmentation may imply that workers and employers in the informal sector, usually associated with small establishments[1], face significant barriers in the transition to the formal sector, with higher productivity and higher income. The second type of segmentation is the inter-regional segmentation, which is mainly due to the barriers of mobility of labor and other factors between regions. Accordingly, the hypothesis of our study is that regional wage inequality may be explained by regional differences in the availability of good jobs that generate higher wages. Meaning that, apart from the differences in the endowment of human capital across Colombian regions, regional heterogeneity in the incidence of informality may be another important source of regional wage disparities.

The empirical analysis consists of examining the returns to education and the pay penalty of informal jobs across Colombians regions by using mean models and quantile regression models in order to analyze the effect of characteristics along the wage distribution. Then, regional wag gaps are decomposed into the contribution of differences in the regional distribution of characteristics, and into the contribution of differences in wage structures (heterogeneity in prices to characteristics). In doing so, we apply the standard Blinder-Oaxaca decomposition at the mean and the decomposition for unconditional quantile regression (UQR) models proposed by Firpo, Fortin and Lemieux (2009, 2011) at selected quantiles. With both of these approaches it is possible to isolate the particular contribution of education and informality to the regional wage gap, in contrast with other procedures (Dinardo, Fortin and Lemieux, 1996; Machado and Mata, 2005; Melly, 2005). Pereira and Galego (2013) applied this method in the case of regional wage differentials for Portugal. As far as we know, our study represents the first application of this method for the analysis of regional wage differentials of a developing country.

Results for Colombia show that regions not only differed in earning relevant characteristics, but also display sizeable regional variability in the returns

---

[1] However, establishment size and sector assignments have been found to be imperfectly correlated.

to these characteristics. Particularly, heterogeneity in returns to education across regions play an important role in explaining regional wage gaps. Additionally, workers face different informal pay penalties throughout the territory and it affects mostly individuals at the lower part of the wage distribution, therefore its contribution in explaining regional wage gaps is limited to this part. Our results confirm previous evidence on the existence of significant regional wage differences between the Golden Triangle region, conformed by the cities of Cali, Medellin and Bogota, and other regions in the country. The difference is particularly wide for those regions with a large share of labor in the informal sector. In fact, after comparing formal workers across regions and separately doing the same for informal workers regional wage gaps are reduced considerably. Furthermore, our results reveal that not distinguishing between formal and informal workers leads to conclusions on the origin of regional wage disparities that are partially misleading. For instance, the belief that the Golden Triangle is the region with the best endowed workforce is not completely accurate when the analysis distinguishes between formal and informal workers. Moreover, it seems that the distribution of education is generating an equalizing effect of wages across some regions, whereas the returns to education continue to be a source of wage inequality across Colombian territories.

The results of this study point to the conclusion that some public policies aim in reducing human capital differences among regions will help to decrease regional wage gaps, especially at the higher parts of the wage distribution. However, equalizing years of education of workers across regions would not be enough to reduce regional wage differences due to the sizeable differences in returns to years of education at higher quantiles. Similar results have been found in previews studies, albeit in a context of developed countries. Meanwhile policies that points towards the reduction of informality will help to minor regional wage gaps at the lower part of the wage distribution particularly for those regions with sizable informality.

The remainder of the paper is organized as follows. The next section presents a description of the data used. Section 3 outline the methodology used in this study. Then, sections 4 and 5 reports and discusses the results. Finally, in section 6 conclusions are presented.

2. Data and descriptive analysis

We use data from the second quarter of 2010 of the Colombian Household Survey (CHS), a repeated cross-section conducted by the National Statistics Department (DANE). The survey gathers information about employment conditions for population aged 12 or more including income, occupation and industry sector at two digit level, in addition to the general population characteristics such as sex, age, marital status and educational attainment. The CHS is representative for the thirteen mayor metropolitan areas in Colombia, composed of a main city and its associated municipalities.

In this study, a sample of 34626 working individuals was drawn from the 2010 CHS. The analysis was restricted to salary workers that were not carrying formal studies aged between 15 and 60 years and who report working more than 16 hours per week. We do not include self-employed and employers workers in the analysis because their source of income is a combination of labor and physical capital and therefore may not be compared with earnings of other employees. Apart from this, self-employed workers' earnings would be expected to have a greater measurement error. Excluding self-employed resulted in dropping 16941 individuals. We also exclude public employees from the sample since by nature they belong to the formal sector and their wages might reflect institutional arrangements. After excluding observations with missing values or inconsistencies for the selected regressors, over 13796 individuals remained in our sample.

The central variable of the analysis is hourly wages. We have combined information from gross monthly income and worked hours in order to obtain gross hourly wages. A fist look at the degree of regional wage differentials in Colombia is obtained from a simple inspection of Table 1, which in the second column displays the average gross hourly wage. Large differences in average wages across the thirteen metropolitan areas are observed. For instance, the average wage in Cucuta, the metropolitan area with the lowest level, was 66.15% of the average wage in Bogotá, the metropolitan area with the highest level. As in previous studies, we attempt to control for price differentials by adjusting the nominal gross hourly wage using the deflator from the consumer price index of each city. Consumer price indices for the main city of each metropolitan area were obtained from DANE. We applied the consumer prices index of the main city to the whole metropolitan area. This implies that the price level of the main city is representative for the whole metropolitan area. The averages of this *adjusted* gross nominal hourly wages are shown in the third column of Table 1. It is observed that the position in the regional ranking of wages is fairly the same and that the metropolitan areas in the top and the bottom of the ranking remain unchanged. The fact that the consumer price index is built with a base year fairly recent, 2008, may explain the small variation obtained after controlling for difference in prices across the metropolitan areas. However, as far as we know this is the only information on relative regional prices available for Colombia.

The regional wage gap observed may be caused because worker´s characteristics differ across the metropolitan areas. In particular, they are known to differ in the workers' endowment of education, which is one of the essential determinants of wages. Table 1 contains the average years of education of workers for each metropolitan area. As it can be seen, there are notable differences in education. On average, workers in Cartagena have more than two years of education than those workers in Cucuta. On the other hand, as has already been mentioned, past studies for Colombia have show that the incidence of informality across regions is remarkably different. Since informal workers earn considerably lower wages than their formal counterparts, then a metropolitan area with a higher proportion of informal workers may have lower wages than a metropolitan area with a low fraction of informal workers. We classify workers as formal or informal according to whether they are covered by the social security system or not. Thus, we define workers as formal if they contribute both to health and old-age insurance. Table 1 also presents the percentage of informal workers in each of the metropolitan areas. In accordance with what has been found in previous studies, the incidence of informality is very different across the metropolitan areas. While Cucuta displays an informality of around 59%, the share of informal workers in Medellin is about 19%. Interestingly, some metropolitan areas with the lower average hourly wages are also those with the highest levels of informality (Villavicencio, Pasto and Cucuta). So these simple descriptive figures suggest a negative correlation between the incidence of informality and the hourly wages in the Colombian metropolitan areas.

In order to make the analysis more tractable and for seek of brevity, metropolitan areas where grouped into regions. In Colombia, six regions have been delimited by geographical proximity and natural characteristics (climate, mountains, proximity to the sea, etc…). According to DANE Colombia is delimited into nine regions: Atlantic, Oriental, Central, Pacific, Bogota, Antioquia, Valle del Cauca, San Andres and Providencia and Orinoquia – Amazonia. Though Bogotá, Antioquia and Valle del Cauca belong to one of the six regions, according to the geographical and natural delimitation, they are taken away from their corresponding region because of their economic importance. In our particular case, we grouped the largest metropolitan areas of these regions (Bogotá, Medellin and Cali, correspondingly) into one region that we will refer as the Golden Triangle[2]. These metropolitan areas are the most dynamic and productive of the country. The most productive firms, most of the R&D

[2] Colombia's Golden Triangle refers to an urban region, limited by a triangle whose vertexes are defined by the three largest cities: Bogotá, Medellin and Cali. In our particular case, we are not referring to the region, but only to the three cities that demarcates the triangle.

investment executed in the country and the highest skill workers are concentrated in these three areas. Although CHS (2010) does not contain information about the metropolitan areas of San Andres and Providencia and Orinoquia-Amazania, there is at least one metropolitan area for each of the remaining regions. Therefore, according to geographical, natural and economic factors we have grouped the metropolitan areas in the dataset into five regions. The first region, Atlantic, includes Barranquilla, Cartagena and Monteria. The second region, Oriental, groups Cucuta, Bucaramanga, and Villavicencio. The third one, Central, it is represented by Manizales, Pereira and Ibague, and the fourth, Pacific, is only composed by Pasto. Finally, the fifth region, Golden Triangle, it is composed by the three largest metropolitan areas of Colombia, Bogotá, Medellin and Cali.

Table 2 provides a description of hourly wages for the five regions. Clearly, average hourly wages differ between regions, although the magnitude of the differences is lower than the one found for the thirteen metropolitan areas. Now, the average hourly wage of the region with the lowest level, Pacific, is 74% of that in the region with the highest level, Golden Triangle. So by grouping metropolitan areas into regions the amount of disparities is attenuated, but they still remain sizable. Apart from the differences in the mean, the wage distributions of these five regions present other interesting variations. For instance, Table 2 shows that the wage distributions of the regions have different degree of dispersion. The standard deviation of the logarithm of gross hourly wages and the Gini index for the region with the lowest level of wages, Pacific, are higher than that of the region with high level of wages, Golden Triangle, suggesting that regions also differ in terms of the amount of intra-regional inequality. Finally, from the value of hourly wages at certain percentiles (25%, 50% and 75%)[3], reported in the last columns of Table 2, it can be concluded that regional wage

[3] In order to save space we do not reproduce here the results in this chapter for other percentiles, although they are available upon request. In any case, including results corresponding to more percentiles does not modify the general conclusions regarding regional disparities over the entire wage distribution.

differentials are far from constant over the entire wage distribution, with symptoms of a no-monotonic behavior.

In order to have a better comparison of the entire wage distributions Figure 1 displays kernel density estimates for hourly wage distributions of the thirteen metropolitan areas and divided into the five regions. Though in particular cases the distribution of hourly wages behaves quite different across the metropolitan areas that comprise each region, in general terms the differences within each region are not notable. In fact, it was expected that some heterogeneity in term of wages and other characteristics remain for some regions, as the grouping criteria not always obeyed to economics factors. On the other hand, Figure 2 displays kernel density estimates for hourly wage distributions once the metropolitan areas are grouped into regions. As it can be seen there are differences in the shape of these distribution. Noticeably, Pacific stands as the region with the higher wage dispersion; its density lies to the left of other regions and displays a higher mass of probability in the lower tail. Oriental and Central regions have a similar pattern as Pacific but less discernible. Hourly wage densities of the Atlantic region and the Golden Triangle are slightly to the right of the rest of other regions and have a narrower left tail. So, the evidence from Table 2 and Figure 3 confirms that there are noticeable differences across regions in the entire wage distribution, and not just on the average wages. To account for these differences, in the rest of this Chapter we provide results for the average and for some selected quantiles.

As has already been mentioned, some of the regional wage differentials might be caused by the spatial distribution of human capital and other earning relevant determinants, as informality. Table 3 reports a simple description of the observable worker and firm characteristics for the five regions. It is for instance observed that regions with high levels of wages have workers employed in relatively larger firms and with a permanent contract. Other differences are worth examining more closely. For example, the proportion of workers employed in the sectors of industry and financial intermediation is larger in high wage regions.

One point that also worth to mentioning is the low proportion of women working in Atlantic region, 39%, compare to 45% in Golden Triangle. Informality also differs considerably between regions; the incidence of informality is 49% in Pacific while in the Golden Triangle is 23%. These differences in the proportion of informal workers across regions might intensify regional wage differentials, since formal jobs usually entail higher wages than informal jobs.

Therefore, there are differences in characteristics between regions that may result in regional wage differentials. Nevertheless the key point is if differences in characteristics can mainly account for regional wage differences, or if part of the wag gap is produced by differences in how these characteristics are paid across regions. If regional wage gap were completely explained by differences in the distribution of observable characteristics across regions, then under such circumstances, similar workers employed in similar firms but located in different regions would earn the same wage. On the contrary, if part of the wage gap could be explained by differences in how characteristics are rewarded, this could be associated to failures in regional labor markets, as similar workers in comparable firms but in different regions would be earning different wages. In the section that follows we aim to shed more light on this issue.

3. Empirical strategy

3.1 Specification of the wage equation

The empirical strategy is based on a model in which the wage of individual i in region r is given by:

$$W_{ir} = X_{ir}\boldsymbol{\beta}_r + \varepsilon_{ir} \quad (1)$$

where W_{ir} denotes the log of the hourly wage of individual i in region r, X_{ir} denotes the set of characteristics that affect the wage of this individual (e.g. education, experience, tenure, sector of employment), and $\boldsymbol{\beta}_r$ is the vector of prices or returns at region r associated to the characteristics in X_{ir}. Estimated returns based on equation (1) using ordinary least square (OLS) are based on the mean of the conditional distribution of wages.

The analysis from equation (1) is based on the mean. However, the descriptive in the previous section showed that regional disparities are far from uniform over the entire wage distribution. Therefore, it is of interest to know the effects of the exogenous variables, for example education, at different points of the distribution of wages. This can be done by using the conditional quantile regression (CQR) model introduced by Koenker and Bassett (1978). It can be written as:

$$W_{ir} = X_{ir}\boldsymbol{\beta}_{\tau r} + \varepsilon_{\tau ir} \quad \text{with} \quad Q_\tau(W_{ir}|X_{ir}) = X_{ir}\boldsymbol{\beta}_{r\tau} \tag{2}$$

where $Q_\tau(W_{ir}|X_{ir})$ denotes the τ-th conditional quantile of wages given the set of characteristics in X_{ir}. Analogous to the OLS regression of W_{ir} on X_{ir}, where $\boldsymbol{\beta}_r$ is estimated as a solution of minimizing sum of square residuals, $\boldsymbol{\beta}_{\tau r}$ associated with τ-th conditional quantile function may be estimated by minimizing a sum of asymmetrically weighted absolute residuals (Koenker, 2005; Koenker and Bassett, 1978):

$$\min_{\beta_{\tau r}} \sum \rho_{\tau r}(W_{ir} - X_{ir}\boldsymbol{\beta}_{\tau r}) \tag{3}$$

where $\rho_{\tau r}$ is the checkpoint function defined as:

$$\rho_{\tau r} = \tau Z \ if \ Z \geq 0$$

or

$$\rho_{\tau r} = (\tau - 1)Z \text{ } if \text{ } Z < 0 \tag{4}$$

The estimated coefficients of $\boldsymbol{\beta}_{\tau r}$ may be interpreted as marginal or partial effects (depending on whether the corresponding covariate is continuous or binary) on the conditional quantile of interest. If $\boldsymbol{\beta}_{\tau r}$ is a consistent estimator of the conditional and unconditional quantile of W_{ir}, then the underlying data generating process follows a linear-in-parameters additive model structure, i.e. is a pure parallel location-shift data generating process for every covariate. However if the conditional effect of a specific exogenous variable in X_{ir} varies over levels of other exogenous variables in X_{ir}, $\boldsymbol{\beta}_{\tau r}$ may be a consistent estimator of the conditional effect of an exogenous variable at the mean values of the other *k-1* remaining exogenous variables, but is not a consistent estimator of the unconditional effect of X_{ir} (Borah and Basu, 2013). Meaning that, for example, the 90th percentile of the unconditional distribution of wages may not be the same as the 90th percentile of wages conditional on years of education.

It is possible to estimate the unconditional quantile effect of X_{ir} using the approach proposed by Firpo, Fortin and Lemieux (2009) based on the influence function (IF) and recentered influence function (RIF). In the context of wages, the IF is:

$$IF(W_{ir}; q_\tau) = (\tau - I\{Y \leq q_\tau\})/f_{W_r}(q_\tau) \tag{5}$$

where q_τ refers to the τ-th unconditional quantile of wages, $f_{W_r}(q_\tau)$ is the probability density function of W_{ir} evaluated at q_τ, and $I\{Y \leq q_\tau\}$ is an indicator variable to denote whether an outcome value is less than q_τ or not. By definition the RIF is equal to:

$$RIF(W_{ir}; q_\tau) = q_\tau + IF(W_{ir}; q_\tau) \tag{6}$$

Firpo, Fortin and Lemieux (2009), demonstrate that the implementation of the UQR is straightforward and similar to the OLS regression. For a specific quantile τ, the first step is to estimate the RIF of the τ-th quantile of W_{ir} following (5) and (6). The second step is to run OLS regression of the $RIF(W_{ir}; q_\tau)$ on the observed covariates, X_{ir}.

$$E[RIF(W_{ir}; q_\tau|\mathrm{X}_{ir})] = X_{ir}\boldsymbol{\beta_{\tau r}} \tag{7}$$

Coefficients $\boldsymbol{\beta}_{\tau r}$ represents the approximate marginal effects of the explanatory variables on the unconditional quantile q_τ of wages for workers in region *r*.

3.2 Decomposition of regional wage gaps

It is possible to obtain a decomposition of the wage differential at quantile τ, similar to the classical Blinder and Oaxaca decomposition, for any two regions using the RIF regression approach by Firpo, Fortin and Lemieux (2009). Any distributional parameter, for example a wage quantile, can be written as a function $q_\tau(F_W)$ of the cumulative distribution of wages, $F_W(W)$. For example the difference in a wage quantile τ, Δ^{q_τ}, between a high wage region (*r*=*h*) and a low wage region (*r*=*l*), can be written as:

$$\begin{aligned}
\Delta^{q_\tau} &= q_\tau\left(F_{w_h|r=h}\right) - q_\tau\left(F_{w_l|r=l}\right) \\
\Delta^{q_\tau} &= \left[q_\tau\left(F_{w_h|r=h}\right) - q_\tau\left(F_{w_l|r=h}\right)\right] + \left[q_\tau\left(F_{w_l|r=h}\right) - q_\tau\left(F_{w_l|r=l}\right)\right] \\
\Delta^{q_\tau} &= \qquad\qquad \Delta_S^{q_\tau} \qquad\qquad + \qquad\qquad \Delta_X^{q_\tau}
\end{aligned} \tag{8}$$

where $\Delta_S^{q_\tau}$ is the wage structure effect and $\Delta_X^{q_\tau}$ is the composition effect. The counterfactual wage quantile $q_\tau\left(F_{w_l|r=h}\right)$ represents the wage quantile that would prevailed if workers observed in the region with high wages, *r*=*h*, had been paid under the wage structure of workers in the low wage region, *r*=*l*. However, as in the case of the Oaxaca-Blinder decomposition for the mean, if the true conditional expectation is not linear, the decomposition based on a linear regression may be biased (Barsky et al., 2002). In the context of decomposition at quantiles, using a reweighted procedure and the RIF-regressions can solve this problem (Firpo, Fortin and Lemieux, 2007, 2011). First a reweighting factor has to be calculated in the following way:

$$\Psi(X) = \frac{\Pr(r=h|\mathrm{X})/\Pr(r=h)}{\Pr(r=l|\mathrm{X})/\Pr(r=l)} \tag{9}$$

Then RIF-regressions are computed for workers in regions l, h and for the counterfactual l^c region, using the weights in $\Psi(X)$, to later calculate the next decomposition:

$$\begin{aligned}\hat{\Delta}^{q_\tau} &= \left[\bar{X}_h\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau h}} - \bar{X}_l^c\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}}\right] + \left[\bar{X}_l^c\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}} - \bar{X}_l\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}\right] \\ \hat{\Delta}^{q_\tau} &= \qquad\quad \hat{\Delta}_S^{q_\tau} \qquad\quad + \qquad\quad \hat{\Delta}_X^{q_\tau}\end{aligned} \tag{10}$$

where $\bar{X}_r, r = l, h$, denote the mean wages in regions l and h, and $\bar{X}_l^c$ is the counterfactual mean for region *l* using the reweighting factor in (9) so to make the distribution of the characteristics in *X* in the region with low wages similar to that of region with high wages.

The wage structure effect can be divided into a pure wage structure effect and a component measuring the reweighting error, as follows:

$$\begin{aligned}\hat{\Delta}_S^{q_\tau} &= \bar{X}_h(\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau h}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}}) + (\bar{X}_h - \bar{X}_l^c)\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}} \\ \hat{\Delta}_S^{q_\tau} &= \hat{\Delta}_{S,p}^{q_\tau} + \hat{\Delta}_{S,e}^{q_\tau}\end{aligned}$$

$$ \tag{11} $$

The reweighting error goes to zero as $\bar{X}_l^c \longrightarrow \bar{X}_h$.

Similarly, the composition effect can be divided into a pure composition effect and a component for the specification error as:

$$\hat{\Delta}_X^{q_\tau} = (\bar{X}_l^c - \bar{X}_l)\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}} + \bar{X}_l \ (\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}}^{\boldsymbol{c}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\tau l}})$$
$$\hat{\Delta}_X^{q_\tau} = \hat{\Delta}_{X,p}^{q_\tau} + \hat{\Delta}_{X,e}^{q_\tau} \tag{12}$$

4. Results

4.1 OLS estimates of the wage equation

Table 4 reports the results of estimating Mincer wage equations by OLS and by quantile regressions (conditional and unconditional) for different quantiles for the five regions. Since the particular focus of this chapter is on the effect of education and informal work, the results are shown only for the estimates of the coefficients associated to years of education and informality, though a large set of controls was included as regressors.[4] The first column in Table 4 contains the estimates in the mean, that is to say the results of the OLS estimates. The estimated returns to schooling for each region are displayed in the upper panel of the table. As expected, there are significant differences in returns to years of education between regions. For example, a higher return to schooling is observed in those regions with the highest levels of wages. The returns to schooling in Atlantic and Golden Triangle are 8.14% and 8.26% respectively. On the other hand, those regions with the lowest levels of hourly wages display the lowest returns to schooling, 5.57% in the Oriental region and 6.82% in Pacific. Thus, in addition to differences in the

[4] The full set of estimates is not reported here to save space, though it is available upon request.

endowment of education, returns to schooling may be thought to be an important factor in explaining wage gaps across regions.

The OLS estimates of the informal pay penalty, reported in the lower panel of Table 4, show a more complex pattern. The Pacific region is the one with the higher pay penalty; informal workers earn 26.8% less than their formal counterparts. However the next region with the higher pay penalty is the Golden Triangle, with a 13.56% pay penalty. Even though the pay penalty is considerably larger in the region with the lower level of wages compare to the region with the highest level of wages, there seems not to be a clear pattern between informal pay penalty and regional wage gap, when comparing for example the Golden Triangle with the Oriental region, since the pay penalty in the last region is lower. Therefore, the OLS results indicate that Colombian region differ not only in the incidence of informality (the share of the informal sector) but also in the difference in wages earned by otherwise similar formal and informal workers.

4.2 Quantile regression estimates of the wage equation

Table 4 also displays the results of estimating the Mincer wage equations by conditional and unconditional quantile regressions. Results concerning conditional quantiles show conditional returns to schooling and pay penalties earnings after adjusting for workers' and firms' characteristics. Information about the dispersion of wages within certain individuals with the same characteristics can be derived from the CQR results. Consistently with previous literature, returns to schooling are heterogeneous and increasing along the quantiles for all regions. CQR results suggest that, in the Golden Triangle region, returns to schooling range from 4.62% for the first quantile to 8.99% for the last quantile of the conditional distribution of wages. While in the Pacific region, returns to schooling in the first quantile are 5.16% and in the last quantile are 7.29%. Interestingly, the returns to schooling are higher for Pacific compare to those in Golden Triangle at lower quantiles, they are fairly the same at the middle part of the distribution, and lower at higher quantiles. The coefficient of years of education increases along the wage distribution for all regions, suggesting that increasing education has an unequalizing effect in the wage distribution.

However interpreting conditional quantile regression results must be done cautiously. A common difficulty associated with interpreting these results is that, as has already been mentioned, the 90th percentile of the unconditional distribution of wages may not be the same as the 90th percentile of the conditional distribution of wages, then the positive and heterogeneous CQR effects do not imply that education has a stronger effect for the rich, but for the conditionally rich, that is, after controlling for all other covariates. The advantage of the UQR approach is that it studies the effects directly on the distribution of income. The UQR results show also a heterogeneous behavior of the returns to schooling along the wage distribution, but it is even more pronounced. Returns to schooling range from 1.18% to 16.17% for the Golden Triangle and from 4.19% to 8.99% in the Pacific region. With the results from UQR now it can be said that the returns to schooling are larger for those individuals located in the upper part of the wage distribution, the rich. As with the CQR, returns to schooling are higher in the Pacific region at lower quantiles, compare to those in Golden Triangle. However, in contrast to what was found with CQR, at the middle part of the distribution, the returns are considerably higher in the Golden Triangle.

Regarding the informal pay penalty. The pay penalty decreases sharply from the lower quantile to the middle quantile and is statistically significant mainly for the lower quantiles. In the case of the Golden Triangle, and accordingly to the CQR results, the pay penalty is of around 16.62% in the lowest quantile and 7.26% in the upper quantile. In Pacific region informal workers faced a penalty of 29.59% in the lowest quantile and 17.37% in the highest. The informal pay penalty at higher quantiles in the case of the UQR turn to be positive for some regions (e.g. Atlantic, Oriental and Central), pointing towards the existence of a premium for informal workers, however such positive coefficients lack of statistical significance. Informality affects negatively mostly those individuals positioned at the lower part of the, conditional and unconditional, wage distribution. The decrease in the pay penalty of informality means that a 1 percentage increase in informal jobs decreases wages more at the bottom than at the top of the wage distribution. In other words a rise in informal jobs will increase wage inequality in all the Colombians regions.

These estimates confirm the positive effect of education on wages and an increasing effect at higher quantiles of the wage distribution. There is substantial regional variability in the returns to schooling. Furthermore, they suggest that difference in returns to years of education may be an important factor explaining wage differentials across regions. On the other hand, workers face different informal pay penalties throughout the territory and it affects mostly individuals at the lower part of the wage distribution, therefore its contribution in explaining regional wage gaps may be limited to this part.

The evidence presented so far confirms that regions not only differed in the endowment of earning relevant characteristics, such as education, but also shows sizeable regional variability in the returns to these characteristics. The next section assesses the contribution of this variability in characteristics and returns to the wage gap across regions.

4.3 Decomposition of regional wage gaps

The decomposition of regional wage differentials in Colombia is analyzed by considering the difference between Golden Triangle, the region with the highest level of wages, and other regions. Estimated regional wage differentials for each region relative to Golden Triangle for the mean and for the selected quantiles are reported in the first row of Table 5. It also contains the global decomposition, in which wage gaps are decomposed in two terms, one that accounts for the contribution attributable to difference in observable characteristics (labeled Total explained by characteristics) and another that corresponds to differences in the wage structure (labeled Total wage structure). Both of these two components can in turn be decomposed in the specific contribution of each factor that determine wages, by using the detail decomposition. Given the main interest in this chapter, the details of the specific contribution of education and informality are presented in the table, while the contributions of the rest of control variables have been grouped in the term labeled *rest*. In addition, results from the decomposition with and without the reweighting are presented in panels A and B respectively.

Wage differentials between regions, calculated at the mean, are all statistically significant. The highest wage gap at the mean is Pacific region, 36%, and the lowest is Atlantic, 9%. Results from the global decomposition without reweighting (Panel A) indicate that the contributions of coefficients are larger than that of characteristics for most of the regions, except for Oriental region. In the case of Atlantic, difference in characteristics pushes down the wage gap, as this region is more endowed than Golden Triangle. However difference in coefficients enlarged the wage gap, meaning that workers characteristics in Golden Triangle are better rewarded than in Atlantic region.

In all regions, except for Atlantic, the specific contribution of education indicates that a considerable part of the wage differential between regions is explained by the fact that the Golden Triangle has a more educated workforce. Golden Triangle also displays the highest returns to schooling, which is reflected in the positive effect in the wage structure. Meanwhile the differences in the incidence of informality across regions suggest that a more equal distribution of informality may reduce the wage gap between regions. In contrast, the difference in the informal pay penalty does not contribute to drive regional wage gaps.

As already discussed, wage differentials at the mean may hide important information of the wage gap across the wage distribution. Table 5 also shows regional wage differentials for each region relative to Golden Triangle at different quantiles. The quantile approach reveals that, for Oriental and Pacific regions, the wage gap along the wage distribution has a non-monotonic behavior. This behavior is different to what has been described for developed countries. Motellón, López-Bazo and El-Attar (2011) found an increasing wage differential across the wage distribution for Spain and Pereira and Gallego (2013) found the same pattern for Portugal.

Regional wage gaps and the decomposition analysis at selected quantiles employing the method in Firpo, Fortin and Lemieux (2009) are also reported in Table 5. For most of the regions and for most of the quantiles differences in coefficients are the dominant effect explaining regional wage gaps. However, Oriental once again stands as the region in which difference in characteristics represents the most part of the wage differential. The specific contribution of education at lower quantiles is not what is driving regional wage differentials. If any, in some cases education pushes down wage differentials at lower quantiles. For example, in the case of Pacific at the 25th quantile difference in returns to schooling reduce the wage gap. However at the middle and at higher quantiles of the wage distribution educations plays an important role and a large part of wage differentials is due to difference in the returns to education. As expected, informality and specifically its incidence only affect regional wage gaps at lower quantiles. Informality represents around 50% of the wage gap at the 25th quantile in Pacific, meaning that reducing informality in this region will help to reduce the wage gap considerably.

With respect to the constant, it is only important in the case of Oriental region. The constant corresponds to the unexplained part, not accounted by covariates. For the other regions is not statistical significant.

Table 5 also displays the decomposition with reweighting and the specification and the reweighting errors. Concerning the reweighting decomposition, one can see that the results change slightly for most regions.

However in the case of Pacific the reweighting decomposition points to a greater contribution of the characteristics component to the wage gap and less to the wage structure, though it remains considerable for the lowest quantile. The specification errors are for some regions and for some quantiles statistical significant and its value is not negligible. As for the reweighting errors, they are quite small for most quantiles and sometimes significant at 5% level. Nevertheless the conclusions derived from the decomposition without reweighting remain fairly the same for most of the regions.

These results lead??? to the conclusion that policies aiming at reducing human capital differences among regions will help to decrease regional wage gaps, especially at the higher parts of the wage distribution. However, equalizing years of education of workers across regions would not be enough to reduce regional wage differences due to the sizeable differences in returns to years of education at higher quantiles. Similar results have been found in previews studies, albeit in a context of developed countries. Meanwhile policies that points towards the reduction of informality will help to lower regional wage gaps at the lower part of the wage distribution particularly for those regions with sizable informality.

5. Regional formal and informal wage gaps

The above results were done jointly for formal and informal workers, thus assuming that returns to education and the effects of other relevant characteristics that determine wages were the same for both type of workers. However, the existence of institutional arraignment or different wage structures may affect the way formal and informal workers are rewarded, and therefore the prices that they perceived for their characteristics. For example, it is well known that the minimum wage is binding in the formal sector, meaning that a large proportion of formal workers earn a minimum wage, in contrast a large proportion of informal workers are paid a wage inferior to the minimum wage. This adds to the fact that the share of workers in the informal sector varies largely across Colombian regions. Thus, grouping formal and informal workers together may gave misleading information about the origin of regional wage disparities.

As shown in Figure 3, once the density of hourly wages is computed for formal and informal workers separately, the regional differences are less marked within each of these two groups of workers. As a matter of fact, Pacific region whose density distribution of hourly wages had a very dissimilar behavior compared to other regions in the total sample, once formal and informal workers are treated separately, its behavior is more alike, especially in the case of formal workers. Table 6 provides a description of hourly wages for formal and informal workers separately and for the five regions, similar to Table 2. Undoubtedly, average hourly wage are different between regions, even after splitting the sample into formal and informal workers. However wage gaps are reduced considerably. By comparing formal workers from Pacific region to formal workers of Golden Triangle now the average of gross hourly wages of Pacific is 95% of that paid in Golden Triangle. When considering only informal workers the wage gap of Pacific against Golden Triangle is also reduced, although to a lesser extent, informal workers at Pacific earn 78% of what informal workers earn at Golden Triangle. The last columns of Table 6 report gross hourly wages at the selected percentiles. The wage gap for formal workers behaves in a different way along the wage distribution for each of the regions. While a no-monotonic behavior throughout the wage distribution is present for informal workers. Since the magnitude and the behavior of regional wage gaps of formal workers are different to those of informal workers, then treating formal and informal workers separately will complement the analysis and will give a more complete??? understanding of regional wage gaps in a labor market characterized by a high degree of informality. In doing so this section will present the same analysis done so far, but differentiating formal and informal workers. However, the focus is not to compare formal and informal workers, but to compare formal workers across regions and separately doing the same for informal workers. While comparing formal workers to their informal counterparts across regions is of interest, is out of the scope of this study. Moreover, the selectivity bias associated with non-observable characteristics that could simultaneously affect wages and the sector in which the individuals are currently working is less likely to affect the results when comparing formal (informal) workers of one region with formal (informal) workers of other region.

Table 7 reports the results concerning the estimates of the Mincer wage equations by OLS and by quantile regressions (conditional and unconditional) for the selected quantiles for the five regions and for formal and informal workers separately. The discussion that follows of this set of results is done taking as a point of reference the results in Table 4 when formal and informal workers were treated jointly with the aim of highlighting the importance of this subsequent analysis. As in the past, the description will focus only on the returns to education. Looking at the results found for formal workers it is observed that returns to education for these type of workers differ across regions but to a less extend than those obtained previously. Results from quantile regression (conditional and unconditional) show that returns to education for formal workers increase along the wage distribution and for specific quantiles some differences between regions exist, but again these difference are lower than those found in Table 4.

Turning now to the results for informal workers it is visible that returns to education differ considerably across regions for this type of workers. The OLS estimates show that informal workers of the Atlantic and Pacific regions have the highest returns, around 5%, while Oriental and, surprisingly, Golden Triangle display the lowest ones, around 3%. The returns to education for informal workers increase along the wage distribution in some cases, such as in Central and Oriental regions, and for other regions they have a non-monotonic behavior. Clearly the results for informal workers differ considerably to those found in Table 4. Moreover they suggest that the value of additional education is quite constraint in the informal sector, as more education not necessary means higher wages.

From these findings is clear that grouping formal and informal workers do not reveal the complete picture and may produce only incomplete conclusions. There are reasons to suspect that the decomposition analysis might also give new information if it is done for formal and informal workers separately. Table 8 and Table 9 display the results of the decomposition exercise for formal and informal workers respectively, similar to the results presented in Table 5 for the entire sample of workers. Results from the global decomposition, for formal and informal workers, show that for Atlantic region the results are fairly the same, but for the rest of the regions the results of the decomposition provide new information. First, it is important to notice that for all regions the characteristics component reduces considerably its contribution. This may be the result of comparing more homogenous workers across regions, especially in the case of formal workers who share similar worker and firms' characteristics. For Pacific region it now turns that formal workers are better endowed that formal workers in the Golden Triangle, and thus the characteristics component reduce the wage gap between this two regions. In the case of Central region the component corresponding to differences in characteristics is not statistical significant neither for formal nor for informal workers. However it remain true that the characteristics of workers in Golden Triangle are better-rewarded compare to other regions.

The detailed decomposition, and particularly the contribution of education, also varies considerably once the analysis is done for formal and informal workers independently. In the case of formal workers, the Atlantic and Pacific regions are endowed with workers with more education than those formal workers in the Golden Triangle, mainly at higher quantiles, and hence just taking into account education the wage gap for these regions will be lower. For central region the difference in education is not statistically significant. Regarding the difference in the returns to education, it can be said that the difference in returns contributed to increase the wage gap.

These results revels that some of the conclusions derived from the preview analysis that treated formal and informal workers jointly are partially correct. For instance the belief that the Golden Triangle is the region with the largest endowed workforce is not completely accurate. Moreover the distribution of education is generating an equalizing effect of wages across some regions. While the returns to education continue to be a source of wage inequality across Colombian territories.

6. Conclusions

Results from micro-data for Colombia confirmed the existence of differences not only in average regional wages but also across the wage distribution. This study used the decomposition approach proposed by Firpo, Fortin and Lemieux (2009) to estimate the contributions of regional differences in characteristics and of regional differences in the wage structures to the observed regional wage gaps. This methodology has the advantage that allows estimating the contribution of each characteristic along the entire wage distribution. Given that Colombian regions are characterized by significant differences in the education of their workforce and in the incidence of informality, the contribution of both of these two factors to the regional wage gaps are closely examined.

The results of the decomposition for Colombia show that for most of the regions and for most of the quantiles differences in the wage structures are the dominant factor explaining regional wage gaps. Meaning that workers with similar characteristics received different wages depending on the region in which they are located. At the middle and especially at higher quantiles of the wage distribution educations plays an important role and a large part of wage differentials is due to differences in the returns to education. Informality and specifically its incidence only affect regional wage gaps at lower quantiles.

Therefore policies that points towards the reduction of informality will help to lower regional wage gaps at the lower part of the wage distribution particularly for those regions with sizable informality.

This study has shown the importance of examining regional wage gaps separately for formal and informal workers since, in addition to the regional disparities in the incidence of informality, it has been proved that the wage structure differ between the two sectors. Accordingly, results on the reasons behind regional wage disparities when distinguishing between workers of the two sectors deviate from those found when they are grouped together. Wage gaps are reduced considerably once formal workers are compared between regions, particularly for those regions with a high incidence of informality. Suggesting that formalization of employment, aside from the well-known implications of higher wages and social security coverage, may also help reducing disparities across regions. Moreover, if regional labor markets are segmented and formal and informal jobs are characterized by different mechanisms of functioning and adjustment, the proposed policy may not be unique for each of these two segments.

As in past studies of this nature, it remains to be explained why the difference in the returns to education across regions is persistent. We hypothesize that such a difference in returns is related to economies of scale and agglomeration economies; however further research is clearly required on this matter for a better understanding of regional wage differentials.

References

Andersson, F., Burgess, S., & Lane, J. I. (2007). Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics*, *61*, 112-128.

Blackaby, D. & Murphy, P. (1995). Earnings, unemployment and Britain's North–South divide: real or imaginary?. *Oxford Bulletin of Economics and Statistics*, 57, 487–512.

Bonet, J., & Meisel, A. (2007). Polarización del ingreso per cápita departamental en Colombia, 1975-2000. *Ensayos Sobre Política Económica*, *25*, 12-43

Bonet J., & Meisel A. (2008). Regional Economic Disparities in Colombia, Investigaciones Regionales, 14, 61–80.

Combes P., Duranton G. & Gobillon L. (2008). Spatial wage disparities: sorting matters!. *Journal of Urban Economics*, *63*, 723–742.

Dinardo J., Fortin N. & Lemieux T. (1996). Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica*, *64*,1001-1044.

Duranton G., & Monastiriotis V. (2002). Mind the gaps: the evolution of regional inequalities in the UK, 1982–1997. *Journal of Regional Science*, *42*, 219–256.

Firpo S., Fortin M., & Lemieux, T. (2009). Unconditional quantile regressions. Econometrica , *77*, 953- 973

Fortin N.M., Lemieux, T., & Firpo S. (2011). Decomposition Methods in Economics, in O. Ashenfelter and D. Card, eds., Handbook of Economics, Amsterdam: North-Holland, Vol. IV.A: 1-102

Galvis, L. (2012). Informalidad Laboral en las Áreas Urbanas de Colombia. *Coyuntura Económica*, *42*, 15-51.

García I., & Molina J. (2002). Inter-regional wage differentials in Spain. *Applied Economic Letters*, *9*, 209–215.

Gasparini, L., Cruces, G. & Tornarolli, L. (2011). Recent Trends in Income Inequality in Latin America. *Economia*, Journal of LACEA.

Glaeser E., Kallal H., Scheinkman J. & Shleifer A. (1992). Growth in cities. *Journal of Political Economy*, *100*, 1126–1152.

Greenwood M. J., Hunt, G. L., Rickman D. S., & Treyz G. I. (1991). Migration, Regional Equilibrium, and the Estimation of Compensating Differentials. *The American Economic Review*, *81*, 1382-1390 .

Joumard, I., & Londoño Vélez, J. (2013). *Income Inequality and Poverty in Colombia - Part 1. The Role of the Labour Market.* OECD Economics Department Working Papers, No. 1036, OECD Publishing.

Koenker, R. (2005). Quantile Regression. Cambridge Books, Cambridge University Press, number 9780521845731, November.

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

López-Bazo, E., & Motellón, E. (2012). Human Capital and Regional Wage Gaps. *Regional Studies*, *46*, 1347-1365.

López-Calva, L. F., & Lustig, N. (eds). (2010). Declining Inequality in Latin America: A decade of progress?. Washington, DC, The Brookings Institution.

Machado J., & Mata J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, *20*, 445-465

Melly B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics* , *12*, 577-590

Motellón E., López-Bazo E. & Attar M. (2011). Regional heterogeneity in wage distributions: evidence from Spain. *Journal of Regional Science*, *51*, 558–584.

Pereira, J. & Galego, A. (2013). Inter-Regional Wage Differentials in Portugal: An Analysis Across the Wage Distribution. *Regional Studies*, DOI:10.1080/00343404.2012.750424

Pereira, J., & Galego, A. (2013). *Decomposition of Regional Wage Differences Along the Wage Distribution in Portugal: the Importance of Covariates*. CEFAGE-UE Working Paper no. 2013-16, University of Évora.

Ortiz, C., Uribe, J., & Badillo, E. (2008). Segmentación inter e intrarregional en el mercado laboral urbano de Colombia, 2001-2006. *Ensayos sobre Política Económica, 27*, 194-231.

Table 1. Hourly wage, informality and human capital variables for the thirteen largest metropolitan of Colombia

	Number of Observations	Nominal Gross Hourly wage (pesos)	Adjusted Hourly wage (pesos)	Schooling (years)	Informality (%)
By metropolitan area					
Barranquilla	1037	3663.16	3510.73	11.31	35.29
		(2947.25)	(2824.61)	(3.45)	
Cartagena	809	3760.54	3605.99	11.74	22.00
		(2518.59)	(2415.08)	(3.44)	
Monteria	759	3650.30	3493.12	11.26	36.89
		(3218.13)	(3079.56)	(3.59)	
Cucuta	754	2825.23	2634.22	9.39	59.15
		(1837.99)	(1713.73)	(4.07)	
Bucaramanga	988	3662.94	3442.25	10.65	31.88
		(2562.04)	(2407.68)	(3.87)	
Villavicencio	862	3306.05	3141.81	10.11	43.85
		(2464.41)	(2341.98)	(3.48)	
Manizales	1109	3506.84	3402.62	11.19	20.83
		(2680.53)	(2600.87)	(3.74)	
Pereira	1014	3351.98	3230.37	10.24	28.60
		(2547.55)	(2455.12)	(3.89)	
Ibague	869	3678.27	3501.31	11.06	36.02
		(2913.20)	(2773.05)	(3.73)	
Pasto	733	2981.61	2885.20	10.53	49.39
		(2668.21)	(2581.93)	(4.14)	
Medellin	1913	3903.84	3718.43	10.96	18.98
		(2904.72)	(2766.76)	(3.76)	
Santafe de Bogota	1754	4305.70	4132.05	11.33	23.95
		(3566.44)	(3422.61)	(3.96)	
Cali	1195	3872.52	3745.43	10.68	28.62
		(3147.60)	(3044.30)	(3.83)	
Colombia	13796	3662.54	3504.48	10.86	31.05
		(2894.79)	(2773.67)	(3.82)	

Notes: Sample means (standard deviation are shown for continuous variables).

Table 2. Descriptive of *Adjusted* Hourly Wages in the Five Regions of Colombia

	Average	Std. Dev. of Logs	Gini	Percentiles 10%	25%	50%	75%	90%
Atlantic	3535.18	0.57	0.33	1631.12	2395.67	2617.42	3727.07	6496.88
Oriental	3108.82	0.54	0.31	1478.28	2000.76	2489.83	3321.36	5188.98
Central	3372.9	0.54	0.32	1635.19	2144.57	2467.86	3489.06	6015.41
Pacific	2885.19	0.69	0.39	940.79	1458.48	2325.62	3010.51	5644.71
Golden Triangle	3874.31	0.57	0.34	1874.24	2384.57	2778.14	4167.22	7165.54
Wage gap								
Atlantic vs. Golden	0.09	-	-	0.13	0.00	0.06	0.11	0.09
Oriental vs. Golden	0.20	-	-	0.21	0.16	0.10	0.20	0.28
Central vs. Golden	0.13	-	-	0.13	0.10	0.11	0.16	0.16
Pacific vs. Golden	0.26	-	-	0.50	0.39	0.16	0.28	0.21

Notes: Sample means. Wage gap = (golden - region$_i$)/ golden.

Table 3. Descriptive of Observable Worker and Firm Characteristics

	Atlantic	Oriental	Central	Pacific	Golden Triangle
Adjusted Hourly Wage	3535.18	3108.82	3372.9	2885.19	3874.28
Informal	0.32	0.44	0.28	0.49	0.23
Worker´s characteristics					
Schooling (years)	11.43	10.1	10.83	10.53	11.03
Experience (years)	18.02	17.09	18.55	17.99	18.05
Tenure (months)	53.91	36.92	48.57	44.74	50.21
Women	0.39	0.43	0.42	0.43	0.45
Married	0.6	0.51	0.49	0.52	0.51
Head of household	0.43	0.4	0.43	0.43	0.44
Type of contract					
No-contract	0.25	0.44	0.26	0.43	0.23
Temporary	0.21	0.21	0.24	0.28	0.24
Permanent	0.54	0.36	0.5	0.29	0.52
Firm size					
Micro	0.27	0.44	0.32	0.5	0.28
Small	0.21	0.21	0.18	0.16	0.2
Medium	0.06	0.05	0.05	0.02	0.07
Large	0.46	0.31	0.44	0.32	0.45
Sector					
Agricultural, mining, electricity, gas and water	0.04	0.04	0.03	0.02	0.02
Industry	0.21	0.19	0.24	0.16	0.26
Construction	0.05	0.11	0.08	0.07	0.06
Sales, Hotels and Restaurants	0.29	0.34	0.26	0.38	0.27
Transportation	0.1	0.08	0.1	0.08	0.07
Financial Intermediation	0.11	0.09	0.11	0.07	0.15
Social Services	0.2	0.16	0.18	0.22	0.17
Observations	*2605*	*2604*	*2992*	*733*	*4862*

Notes: Sample means.

Table 4. Estimations of Wage Equations for the Five Regions of Colombia – OLS and Quantile estimates

		CQR			UQR		
	OLS	25	50	75	25	50	75
Years of education							
Atlantic	0.0826**	0.0553**	0.0697**	0.0873**	0.0087**	0.0435**	0.1319**
	[0.0028]	[0.0020]	[0.0029]	[0.0035]	[0.0012]	[0.0025]	[0.0056]
Oriental	0.0557**	0.0353**	0.0440**	0.0557**	0.0215**	0.0253**	0.0740**
	[0.0027]	[0.0024]	[0.0025]	[0.0035]	[0.0036]	[0.0022]	[0.0046]
Central	0.0752**	0.0412**	0.0569**	0.0779**	0.0214**	0.0306**	0.1148**
	[0.0024]	[0.0016]	[0.0023]	[0.0034]	[0.0024]	[0.0016]	[0.0048]
Pacific	0.0682**	0.0516**	0.0659**	0.0729**	0.0419**	0.0288**	0.0899**
	[0.0050]	[0.0051]	[0.0053]	[0.0083]	[0.0099]	[0.0051]	[0.0079]
Golden Triangle	0.0814**	0.0462**	0.0674**	0.0899**	0.0118**	0.0519**	0.1617**
	[0.0020]	[0.0014]	[0.0017]	[0.0032]	[0.0011]	[0.0019]	[0.0047]
Colombia	0.0742**	0.0460**	0.0597**	0.0778**	0.0139**	0.0374**	0.1254**
	[0.0012]	[0.0009]	[0.0011]	[0.0020]	[0.0009]	[0.0010]	[0.0024]
Informality							
Atlantic	-0.1023**	-0.1691**	-0.0435+	-0.0475+	-0.1137**	-0.0874**	-0.0472
	[0.0257]	[0.0192]	[0.0264]	[0.0265]	[0.0138]	[0.0258]	[0.0525]
Oriental	-0.0991**	-0.1341**	-0.0516*	-0.0599*	-0.2710**	-0.0810**	0.0123
	[0.0257]	[0.0224]	[0.0234]	[0.0302]	[0.0355]	[0.0231]	[0.0445]
Central	-0.0951**	-0.1704**	-0.0515*	0.0159	-0.2389**	-0.0572**	0.0414
	[0.0274]	[0.0183]	[0.0263]	[0.0341]	[0.0326]	[0.0215]	[0.0493]
Pacific	-0.2680**	-0.2959**	-0.2422**	-0.1737+	-0.3085*	-0.3499**	-0.2939**
	[0.0558]	[0.0573]	[0.0595]	[0.0893]	[0.1200]	[0.0642]	[0.0868]
Golden Triangle	-0.1356**	-0.1662**	-0.1091**	-0.0726*	-0.1473**	-0.0470+	-0.0215
	[0.0227]	[0.0169]	[0.0195]	[0.0298]	[0.0147]	[0.0249]	[0.0487]
Colombia	-0.1430**	-0.1927**	-0.0891**	-0.0856**	-0.1881**	-0.0917**	-0.0471+
	[0.0125]	[0.0096]	[0.0116]	[0.0186]	[0.0109]	[0.0118]	[0.0242]

Notes: experience (and its square), tenure (and its square), marital status, head of household, hours worked, type of contract, size of the firm and firm sector are included as controls.

Standard errors in [].+ p<0.1, * p<0.05, ** p<0.01.

Table 5. Regional Wage Gap Decomposition

Atlantic	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.087 ****	**0.006**	**0.068 ****	**0.114 ****	**0.087 ****	**0.006**	**0.068 ****	**0.114 ****
Composition Effect attributable to								
Education	-0.033 **	-0.005 **	-0.021 **	-0.065 **	-0.020 **	-0.002 *	-0.009 *	-0.033 *
Informality	0.012 **	0.013 **	0.004 +	0.002	0.012 **	0.013 **	0.007 **	-0.006
Rest	-0.027 **	-0.011 **	-0.024 **	-0.050 **	-0.023 **	-0.008 **	-0.013 **	-0.017 +
Error					0.010	-0.002	-0.007	0.027
Total explained by characteristics	**-0.049 ****	**-0.003**	**-0.041**	**-0.114 ****	**-0.021 ****	**0.002**	**-0.022 ***	**-0.030**
Wage structure effects attributable to								
Education	-0.014	0.035 +	0.096 *	0.340 **	0.052	0.048 **	0.190 **	0.378 **
Informality	-0.011	-0.011 *	0.013	0.008	0.000	0.001	0.009	-0.022
Rest	0.110	-0.093 *	0.074	0.003	0.137	0.004	0.0345	-0.200
Constant	0.051	0.077 *	-0.073	-0.124	-0.063	-0.044	-0.130	0.025
Error					-0.018 +	-0.005	-0.013 *	-0.037 *
Total wage structure	**0.136 ****	**0.009 +**	**0.110 ****	**0.228 ****	**0.109 ****	**0.004**	**0.091 ****	**0.144 ****

Oriental	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.190 ****	**0.187 ****	**0.118 ****	**0.238 ****	**0.190 ****	**0.187 ****	**0.118 ****	**0.238 ****
Composition Effect attributable to								
Education	0.075 **	0.011 **	0.048 **	0.149 **	0.067 **	0.014 **	0.045 **	0.113 **
Informality	0.028 **	0.030 **	0.010 *	0.004	0.017 **	0.035 **	0.004	-0.018
Rest	0.063 **	0.044 **	0.053 **	0.075 **	0.090 **	0.057 **	0.075 **	0.114 **
Error					-0.004	0.058 **	-0.008	0.013
Total explained by characteristics	**0.166 ****	**0.086 ****	**0.111 ****	**0.228 ****	**0.171 ****	**0.165 ****	**0.115 ****	**0.221 ****
Wage structure effects attributable to								
Education	0.260 **	-0.098 **	0.270 **	0.885 **	0.100 **	-0.037 **	0.046	0.450 **
Informality	-0.016	0.054 **	0.015	-0.015	-0.012	0.005	-0.006	-0.026
Rest	0.041	0.389 **	0.0727	-0.137	0.139	0.174 **	0.118	-0.211
Constant	-0.260 **	-0.244 *	-0.350 **	-0.724 **	-0.203 +	-0.118	-0.151	-0.188
Error					-0.004	0.024	0.007	-0.009
Total wage structure	**0.025**	**0.101 ****	**0.007**	**0.010**	**0.019**	**0.022**	**0.003**	**0.017**

Central	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.119 ****	**0.111 ****	**0.127 ****	**0.189 ****	**0.119 ****	**0.111 ****	**0.127 ****	**0.189 ****
Composition Effect attributable to								
Education	0.016 *	0.002 *	0.010 *	0.031 *	0.021 **	0.005 **	0.010 **	0.035 **
Informality	0.006 **	0.007 **	0.002 +	0.001	0.005 **	0.010 **	0.003 *	-0.004
Rest	0.007	0.009 **	0.007	-0.001	0.007	0.011 *	0.006	0.007
Error					0.003	0.017 *	0.021 **	0.012
Total explained by characteristics	**0.029 ****	**0.018 ****	**0.019 ****	**0.031 ****	**0.036 ****	**0.043 ****	**0.040 ****	**0.051 ****
Wage structure effects attributable to								
Education	0.067 *	-0.104 **	0.232 **	0.507 **	0.029	-0.061 **	0.152 **	0.339 **
Informality	-0.011	0.026 **	0.003	-0.018	-0.005	0.014 *	0.004	-0.025
Rest	0.016	0.321 **	-0.029	-0.077	0.025	0.222 **	0.005	-0.069
Constant	0.018	-0.149 +	-0.098	-0.255	0.042	-0.104	-0.069	-0.093
Error					-0.007	-0.002	-0.005	-0.014
Total wage structure	**0.090 ****	**0.093 ****	**0.108 ****	**0.157 ****	**0.083 ****	**0.069 ****	**0.087 ****	**0.138 ****

Notes: + p<0.1, * p<0.05, ** p<0.01.

Table 5 continue

Pacific	A. Without reweighting								B. With reweighting							
			Quantiles								Quantiles					
	OLS		25		50		75		OLS		25		50		75	
Overall wage gap	**0.362**	**	**0.499**	**	**0.180**	**	**0.334**	**	**0.362**	**	**0.499**	**	**0.180**	**	**0.334**	**
Composition Effect attributable to																
Education	0.040	**	0.006	**	0.026	**	0.080	**	0.065	**	0.023	**	0.040	**	0.105	**
Informality	0.036	**	0.039	**	0.012	+	0.006		0.066	**	0.165	**	0.049	**	0.022	
Rest	0.033	**	0.018	**	0.025	**	0.042	*	0.078	**	0.098	**	0.049	*	0.075	*
Error									-0.002		0.018		-0.091	**	0.080	+
Total explained by characteristics	**0.108**	**	**0.062**	**	**0.063**	**	**0.127**	**	**0.207**	**	**0.303**	**	**0.047**	+	**0.283**	**
Wage structure effects attributable to																
Education	0.139	**	-0.317	**	0.243	**	0.755	**	0.068		-0.169	*	0.054		0.443	**
Informality	0.065	*	0.080		0.150	**	0.135	**	0.028	+	0.117	**	0.034	*	0.016	
Rest	-0.140		-0.121		0.011		-0.401		-0.093		0.319		-0.099		-0.231	
Constant	0.189		0.796	*	-0.287		-0.283		0.179		-0.063		0.164		-0.123	
Error									-0.028	+	-0.008		-0.021	+	-0.054	+
Total wage structure	**0.253**	**	**0.437**	**	**0.117**	**	**0.207**	**	**0.155**	**	**0.196**	**	**0.132**		**0.051**	

Notes: + p<0.1, * p<0.05, ** p<0.01.

Table 6. Descriptive of Hourly Wages for formal and informal workers

Formal

	Average	Std. Dev. of Logs	Gini	Percentiles 10%	25%	50%	75%	90%
Atlantic	4070.65	0.509	0.31	2395.67	2400.59	2888.48	4195.21	7442.87
Oriental	3805.30	0.508	0.30	2039.23	2352.64	2838.29	4111.41	6834.07
Central	3812.80	0.501	0.31	2078.56	2412.64	2793.18	3967.82	6791.97
Pacific	4101.96	0.553	0.34	1966.44	2422.52	2822.36	4515.77	8429.44
Golden Triangle	4292.18	0.548	0.34	2289.19	2402.51	2985.64	4665.06	8260.35
Wage gap								
Atlantic vs. Golden	0.06	-	-	-0.06	0.00	0.03	0.11	0.11
Oriental vs. Golden	0.13	-	-	0.13	0.02	0.05	0.13	0.20
Central vs. Golden	0.12	-	-	0.11	0.00	0.07	0.17	0.20
Pacific vs. Golden	0.05	-	-	0.17	-0.01	0.06	0.04	-0.02

Informal

	Average	Std. Dev. of Logs	Gini	Percentiles 10%	25%	50%	75%	90%
Atlantic	2377.82	0.53	0.29	1063.266	1594.9	2232.859	2608.945	3577.982
Oriental	2213.00	0.44	0.23	1208.654	1582.239	2105.398	2610.694	3289.126
Central	2234.65	0.47	0.25	1181.891	1572.214	2056.551	2498.527	3223.331
Pacific	1638.18	0.51	0.28	711.6368	1023.022	1477.888	2037.228	2533.064
Golden Triangle	2486.12	0.48	0.26	1267.638	1751.266	2256.756	2799.037	3776.234
Wage gap								
Atlantic vs. Golden	0.	-	-	0.16	0.09	0.01	0.07	0.05
Oriental vs. Golden	0.07	-	-	0.05	0.10	0.07	0.07	0.13
Central vs. Golden	0.06	-	-	0.07	0.10	0.09	0.11	0.15
Pacific vs. Golden	0.22	-	-	0.44	0.42	0.35	0.27	0.33

Notes: Sample means. Wage gap = (golden - region $_i$)/ golden.

Table 7. Estimations of Wage Equations for the Five Regions of Colombia for Formal and Informal workers – OLS and Quantile estimates

		CQR			UQR		
	OLS	25	50	75	25	50	75
Formal							
Atlantic	0.0955**	0.0554**	0.0845**	0.1041**	0.0218**	0.0769**	0.1693**
	[0.0034]	[0.0020]	[0.0031]	[0.0051]	[0.0022]	[0.0037]	[0.0079]
Oriental	0.0740**	0.0371**	0.0623**	0.0845**	0.0157**	0.0509**	0.1206**
	[0.0036]	[0.0026]	[0.0033]	[0.0070]	[0.0025]	[0.0037]	[0.0074]
Central	0.0879**	0.0371**	0.0753**	0.0952**	0.0127**	0.0644**	0.1803**
	[0.0028]	[0.0013]	[0.0029]	[0.0061]	[0.0016]	[0.0032]	[0.0069]
Pacific	0.0824**	0.0576**	0.0799**	0.0980**	0.0163**	0.0718**	0.1605**
	[0.0070]	[0.0039]	[0.0052]	[0.0148]	[0.0053]	[0.0083]	[0.0170]
Golden Triangle	0.0950**	0.0515**	0.0808**	0.1062**	0.0182**	0.0777**	0.1699**
	[0.0022]	[0.0015]	[0.0025]	[0.0042]	[0.0014]	[0.0025]	[0.0049]
Colombia	0.0890**	0.0449**	0.0756**	0.0975**	0.0159**	0.0707**	0.1728**
	[0.0014]	[0.0007]	[0.0018]	[0.0027]	[0.0008]	[0.0016]	[0.0034]
Informal							
Atlantic	0.0551**	0.0480**	0.0409**	0.0429**	0.0340**	0.0231**	0.0330**
	[0.0051]	[0.0083]	[0.0066]	[0.0052]	[0.0077]	[0.0042]	[0.0050]
Oriental	0.0289**	0.0258**	0.0266**	0.0251**	0.0178**	0.0284**	0.0274**
	[0.0041]	[0.0059]	[0.0035]	[0.0034]	[0.0059]	[0.0045]	[0.0045]
Central	0.0486**	0.0374**	0.0349**	0.0401**	0.0288**	0.0295**	0.0363**
	[0.0046]	[0.0054]	[0.0040]	[0.0044]	[0.0057]	[0.0044]	[0.0045]
Pacific	0.0508**	0.0497**	0.0543**	0.0451**	0.0290**	0.0522**	0.0471**
	[0.0069]	[0.0097]	[0.0102]	[0.0089]	[0.0109]	[0.0098]	[0.0098]
Golden Triangle	0.0346**	0.0277**	0.0226**	0.0291**	0.0186**	0.0183**	0.0250**
	[0.0044]	[0.0057]	[0.0040]	[0.0046]	[0.0052]	[0.0038]	[0.0048]
Colombia	0.0415**	0.0344**	0.0322**	0.0324**	0.0255**	0.0287**	0.0309**
	[0.0022]	[0.0030]	[0.0019]	[0.0020]	[0.0032]	[0.0024]	[0.0023]

Notes: experience (and its square), tenure (and its square), marital status, head of household, hours worked, type of contract, size of the firm and sector are included as controls.

Standard errors in [].+ $p<0.1$, * $p<0.05$, ** $p<0.01$.

Table 8. Regional Wage Gap Decomposition Formal Workers

Atlantic	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.032 ***	**0.015 +**	**0.032 ***	**0.099 ****	**0.032 ****	**0.015 +**	**0.032 ***	**0.099 ****
Composition Effect attributable to								
Education	-0.055 **	-0.011 **	-0.045 **	-0.098 **	-0.040 **	-0.004 **	-0.030 **	-0.067 **
Rest	-0.043 **	-0.022 **	-0.041 **	-0.062 **	-0.037 **	-0.016 **	-0.027 **	-0.033 **
Error					0.011	0.054 **	0.044 **	-0.006
Total explained by characteristics	**-0.098 ****	**-0.033 ****	**-0.086 ****	**-0.160 ****	**-0.066 ****	**0.034 ****	**-0.014**	**-0.106 ****
Wage structure effects attributable to								
Education	-0.007	-0.044	0.010	0.007	0.161 **	0.114 **	0.201 **	0.398 **
Rest	0.070	0.074	0.065	0.008	0.033	0.119 +	0.211	-0.272
Constant	0.067	0.018	0.043	0.244	-0.077	-0.247 **	-0.351 *	0.116
Error					-0.019 +	-0.005	-0.015	-0.037 *
Total wage structure	**0.130 ****	**0.048 ****	**0.118 ****	**0.259 ****	**0.098 ****	**-0.020 ****	**0.046 ****	**0.205 ****

Oriental	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.009 ****	**0.010**	**0.053 ****	**0.130 ****	**0.090 ****	**0.010**	**0.053 ****	**0.130 ****
Composition Effect attributable to								
Education	0.024 *	0.005 *	0.020 *	0.043 *	0.019 +	0.006 +	0.017 +	0.034 +
Rest	0.035 **	0.015 **	0.033 **	0.044 **	0.050 **	0.021 **	0.045 **	0.076 **
Error						-0.035 **	0.002	-0.007
Total explained by characteristics	**0.059 ****	**0.020 ****	**0.053 ****	**0.087 ****	**0.069 ****	**-0.007**	**0.063 ****	**0.104 ****
Wage structure effects attributable to								
Education	0.239 **	0.028	0.305 **	0.561 **	0.152 **	-0.074 +	0.084	0.289 **
Rest	0.035	0.207 **	0.093	-0.217	0.092	0.192 +	0.080	-0.129
Constant	-0.243 +	-0.244 **	-0.398 **	-0.301	-0.219	-0.098	-0.170	-0.124
Error					-0.005	-0.003	-0.005	-0.009
Total wage structure	**0.031 ***	**-0.010**	**0.000**	**0.043 +**	**0.021**	**0.017**	**-0.010**	**0.027**

Central	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.093 ****	**0.003**	**0.069 ****	**0.169 ****	**0.093 ****	**0.003**	**0.069 ****	**0.169 ****
Composition Effect attributable to								
Education	-0.006	-0.001	-0.005	-0.011	-0.001	0.000	-0.001	-0.002
Rest	0.006	0.004	0.005	0.006	0.008	0.006	0.008	0.007
Error					0.007	0.001	0.007	0.016
Total explained by characteristics	**0.000**	**0.003**	**0.000**	**-0.005**	**0.014**	**0.007**	**0.014**	**0.021**
Wage structure effects attributable to								
Education	0.083 *	0.064 *	0.156 **	-0.122	0.094 *	0.059 *	0.114 *	-0.162
Rest	0.100	0.137 *	0.245 +	0.208	0.067	0.120 +	0.226 +	0.188
Constant	-0.089	-0.200 **	-0.332 *	0.089	-0.074	-0.182 *	-0.278 +	0.136
Error					-0.007	-0.002	-0.006	-0.012
Total wage structure	**0.094 ****	**0.000**	**0.069 ****	**0.175 ****	**0.079 ****	**-0.004**	**0.055 ****	**0.149 ****

Notes: + p<0.1, * p<0.05, ** p<0.01.

Table 8 continue

Pacific	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.043**	**-0.025**	**0.055 +**	**0.018**	**0.043**	**-0.025**	**0.055 +**	**0.018 0**
Composition Effect attributable to								
Education	-0.058 **	-0.011 **	-0.048 **	-0.104 **	-0.018	-0.005	-0.020	-0.024
Rest	0.001	-0.012 +	-0.005	0.014	-0.033	-0.011	0.028	-0.006
Error					-0.008	0.004	-0.026	-0.038
Total explained by characteristics	**-0.057 ***	**-0.023 ****	**-0.053 ****	**-0.089 ***	**-0.060**	**-0.012**	**-0.018**	**-0.069**
Wage structure effects attributable to								
Education	0.154 +	0.022	0.073	0.114	0.231 **	-0.020	-0.059	0.820 **
Rest	0.034	0.405 *	-0.027	-0.099	-0.034	0.394 *	-0.065	-0.436
Constant	-0.088	-0.429 *	0.061	0.092	-0.070	-0.382 *	0.218	-0.250
Error					-0.026	-0.006	-0.021	-0.048
Total wage structure	**0.100 ****	**-0.002**	**0.108 ****	**0.107 ***	**0.102 ****	**-0.013**	**0.073 ***	**0.087 ***

Notes: + p<0.1, * p<0.05, ** p<0.01.

Table 9. Regional Wage Gap Decomposition Informal Workers

Atlantic	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.077 ****	**0.094 ****	**0.011**	**0.083 ****	**0.077 ****	**0.094 ****	**0.011**	**0.083 ****
Composition Effect attributable to								
Education	-0.025 **	-0.014 **	-0.013 **	-0.018 **	-0.036 **	-0.036 **	-0.022 **	-0.021 **
Rest	-0.061 **	-0.051 **	-0.039 **	-0.059 **	-0.061 **	-0.070 *	-0.061 **	-0.050 **
Error					-0.001	-0.011	-0.083 **	0.030
Total explained by characteristics	**-0.086 ****	**-0.065 ****	**-0.052 ****	**-0.077 ****	**-0.098 ****	**-0.117 ****	**-0.165 ****	**-0.041 +**
Wage structure effects attributable to								
Education	-0.199 **	-0.151	-0.046	-0.078	-0.144 *	-0.291 **	-0.115 +	-0.042
Rest	0.518 **	0.757 **	-0.025	0.116	0.474 *	1.349 **	0.048	-0.040
Constant	-0.155	-0.448	0.134	0.122	-0.153	-0.851 *	0.241	0.207
Error					-0.002	0.003	0.001	-0.001
Total wage structure	**0.164 ****	**0.159 ****	**0.063 ****	**0.160 ****	**0.175 ****	**0.211 ****	**0.176 ****	**0.124 ****

Oriental	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.093 ****	**0.103 ****	**0.082 ****	**0.074 ****	**0.093 ****	**0.103 ****	**0.082 ****	**0.074 ****
Composition Effect attributable to								
Education	0.019 **	0.010 *	0.010 **	0.014 **	0.017 **	0.012 *	0.014 **	0.020 **
Rest	0.035 **	0.038 **	0.032 **	0.036 **	0.035 **	0.022 +	0.027 *	0.054 **
Error					-0.001	-0.013	0.002	-0.014
Total explained by characteristics	**0.054 ****	**0.049 ****	**0.042 ****	**0.050 ****	**0.052 ****	**0.021**	**0.043 ***	**0.059 ****
Wage structure effects attributable to								
Education	0.048	0.007	-0.085 +	-0.020	-0.010	-0.064	-0.100 +	-0.141 *
Rest	0.285	0.474 *	0.290 +	0.204	0.235	0.487 +	0.277	0.311
Constant	-0.294	-0.426 +	-0.165	-0.160	-0.186	-0.344	-0.140	-0.156
Error					0.002	0.003	0.003	0.001
Total wage structure	**0.039 ***	**0.054 ***	**0.040 ***	**0.024**	**0.041 ***	**0.082 ****	**0.039 ***	**0.015**

Central	A. Without reweighting				B. With reweighting			
		Quantiles				Quantiles		
	OLS	25	50	75	OLS	25	50	75
Overall wage gap	**0.105 ****	**0.105 ****	**0.105 ****	**0.119 ****	**0.105 ****	**0.105 ****	**0.105 ****	**0.119 ****
Composition Effect attributable to								
Education	0.014 *	0.008 *	0.008 *	0.010 *	0.023 *	0.014 *	0.015 *	0.018 *
Rest	-0.008	0.008	-0.004	-0.008	-0.029 *	-0.021 +	-0.025 *	-0.026 *
Error					-0.001	-0.005	-0.006	0.002
Total explained by characteristics	**0.006**	**0.015**	**0.004**	**0.002**	**-0.007**	**-0.012**	**-0.017**	**-0.007**
Wage structure effects attributable to								
Education	-0.121 *	-0.088	-0.096 +	-0.097 +	-0.144 *	-0.102	-0.119 *	-0.120 *
Rest	0.044	0.236	0.169	-0.059	-0.009	0.127	0.083	-0.073
Constant	0.175	-0.059	0.029	0.272	0.266	0.087	0.154	0.316
Error					-0.001	0.005	0.002	0.002
Total wage structure	**0.098 ****	**0.090 ****	**0.101 ****	**0.116 ****	**0.112 ****	**0.117 ****	**0.122 ****	**0.126 ****

Notes: + p<0.1, * p<0.05, ** p<0.01.

Table 9 continue

Pacific	A. Without reweighting								B. With reweighting							
			Quantiles								Quantiles					
	OLS		25		50		75		OLS		25		50		75	
Overall wage gap	**0.433**	**	**0.535**	**	**0.433**	**	**0.333**	**	**0.433**	**	**0.535**	**	**0.433**	**	**0.333**	**
Composition Effect attributable to																
Education	0.009		0.005		0.005		0.006		-0.003		-0.002		-0.003		-0.002	
Rest	-0.022		-0.012		-0.017		-0.025	+	-0.032		-0.043		-0.042		-0.031	
Error									0.013		0.001		0.049		0.019	
Total explained by characteristics	**-0.014**		**-0.008**		**-0.012**		**-0.019**		**-0.022**		**-0.044**		**0.005**		**-0.014**	
Wage structure effects attributable to																
Education	-0.142	*	-0.092		-0.297	**	-0.194	*	-0.181	*	-0.133		-0.317	**	-0.208	*
Rest	0.285		0.503		0.602		-0.179		0.241		0.455		0.393		-0.264	
Constant	0.305		0.131		0.141		0.725		0.393		0.256		0.351		0.820	+
Error									0.002		0.001		0.001		0.000	
Total wage structure	**0.447**	**	**0.543**	**	**0.445**	**	**0.353**	**	**0.455**	**	**0.579**	**	**0.427**	**	**0.348**	**

Notes: + p<0.1, * p<0.05, ** p<0.01.

Figure 1. Regional hourly wage kernel density estimates - Thirteen largest metropolitan areas of Colombia

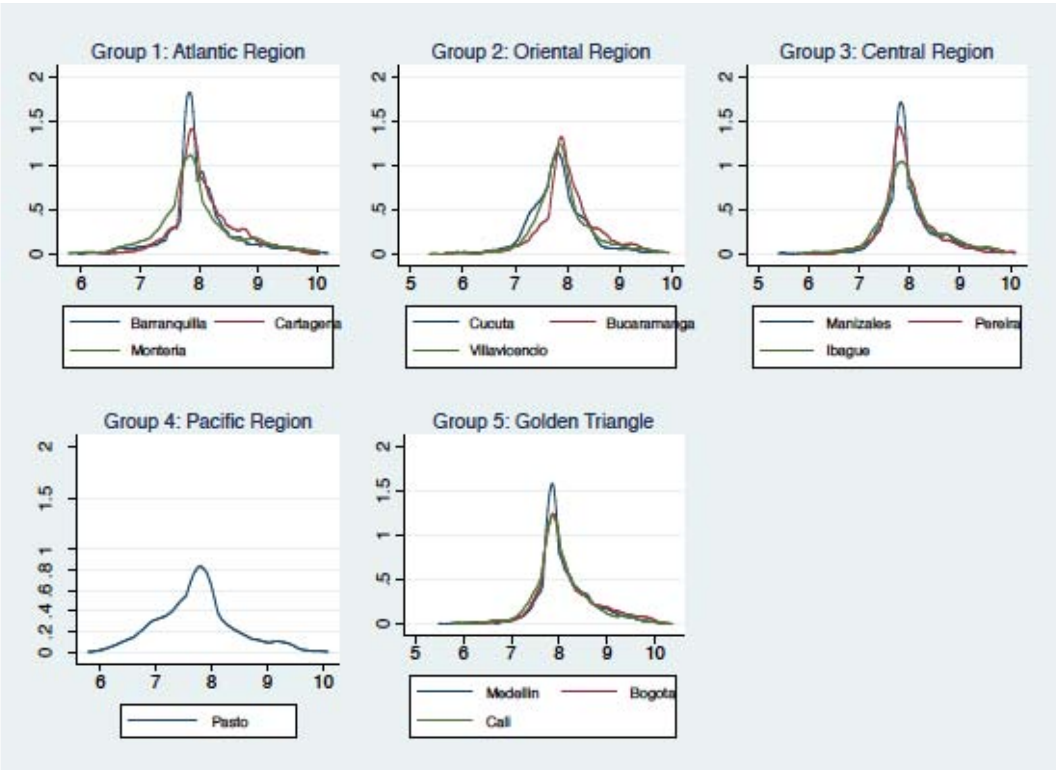

Figure 2. Regional hourly wage kernel density estimates - Five regions of Colombia

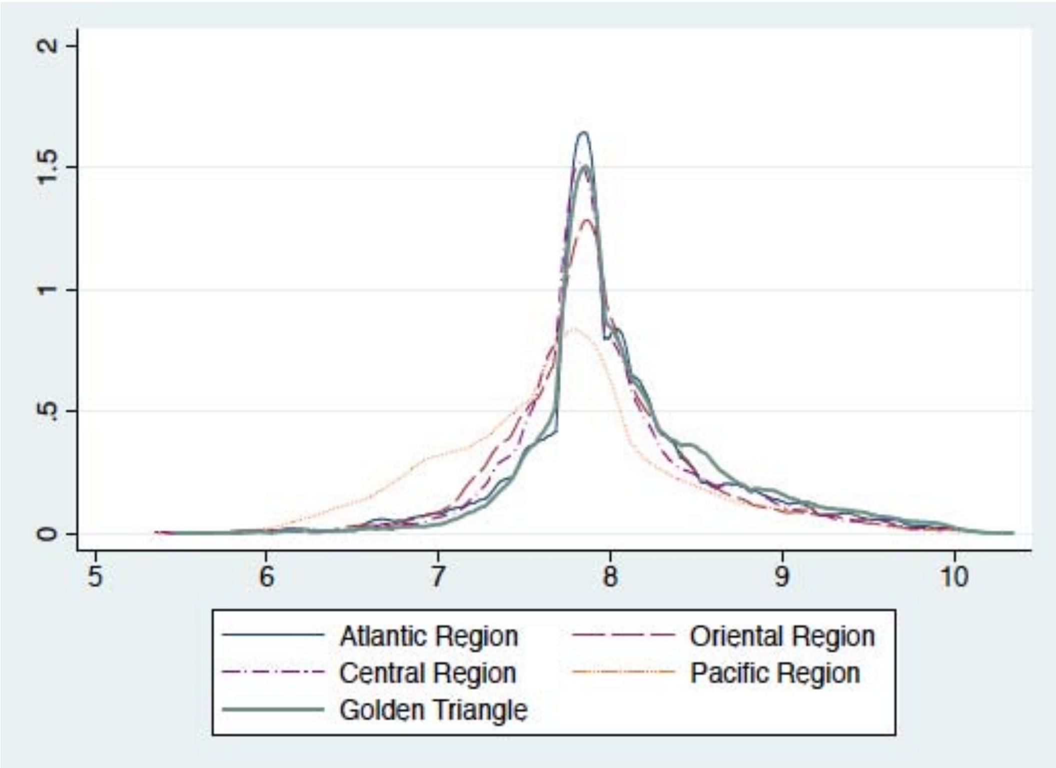

Figure 3. Formal and Informal hourly wage kernel density estimates - Five regions of Colombia