IS THE WAGE CURVE FORMAL OR INFORMAL?
EVIDENCE FOR COLOMBIA[1]

Raul Ramos[a],*, Juan C. Duque[b], Jordi Surinach[a]

[a] Grup d'Anàlisi Quantitativa Regional (AQR-IREA), Department of Econometrics, Statistics and Spanish Economy, Universitat de Barcelona, Avda. Diagonal, 690, 08034 Barcelona, Spain.

[b] Research in Spatial Economics (RiSE-group), Department of Economics, EAFIT University. Carrera 49 N. 7 Sur – 50, Medellin, Colombia.

Abstract: The objective of this paper is to analyse the existence or not of a wage curve in Colombia, paying special attention to the differences between formal and informal workers, an issue that has been systematically ignored in the wage curve literature.

The obtained results using microdata from the Colombian Continuous Household Survey (CHS) between 2002 and 2006 show the existence of a wage curve with a negative slope for the Colombian economy. Using information on metropolitan areas, the estimates of the elasticity of individual wages to local unemployment rates was -0.07, a value that is very close to those obtained for other countries. However, the disaggregation of statistical information for formal and informal workers has shown significant differences among both groups of workers. In particular, for the less protected groups of the labour market, informal workers (both men and women), a high negatively sloped wage curve was found. This result is consistent with the conclusions from efficiency wage theoretical models and should be taken into account when analysing the functioning of regional labour markets in developing countries.

Keywords: Wage curve, unemployment, formal and informal sectors

JEL classification: J30, J60, O17

1. Motivation and objectives

According to a recent OECD study, informal employment account for more than the 50% of total employment in Latin America (OECD, 2009). Colombia is one of the countries in the area with a lower share of informal workers, although it is around the 40% of total workers. Informal employment can be the result of both, people being excluded from formal jobs, or people voluntary opting out of formal structures to avoid paying for taxes and social contributions. However, informal employment usually implies that people are trapped in unproductive, precarious and less protected jobs that will also be much more exposed to local labour market conditions than jobs in the formal sector.

Since the contribution by Blanchflower and Oswald (1990), several studies have focused on the analysis of the relationship between individual wages and local unemployment. For different countries and time periods, this literature has found a negative and significant relationship between individual wages and local unemployment, with a value of the elasticity close to -0.1.[2] Although the list of considered countries is huge, there are only some studies that have focused on Latin America,[3] and only three of them have considered the labour market duality between formal and informal workers, which is clearly relevant in this geographical area and has been totally ignored in the studies for developed economies. In particular, Berg and Contreras (2004) found no evidence of significant differences in the response of wages to local unemployment in the case of Chile while Castro (2006) for Mexico and Bucheli and González (2007) for Uruguay found higher values of the elasticity for informal workers.

Regarding Colombia, the only work on wage curve (Sánchez and Nuñez, 1998) does not provide any information about whether or not there exist differences among formal and informal workers. In their study, the authors found that during the period 1984-1996, the elasticity between wages and unemployment is -0.13. Taking this into account, the main objective of this paper is to provide new evidence on the existence of a wage curve in Colombia, paying special attention to the differences between formal and informal workers.

---

[2] See Blanchflower and Oswald (2005), Nijkamp and Poot (2005) or Babécky et al. (2008) for recent surveys.
[3] Barros and Mendonca (1996) tested the existence of a wage curve in Brazil; Sánchez and Nuñez (1998) provided estimations for Colombia; Galiani (1999) for Argentina; Berg and Contreras (2004) for Chile; Castro (2006) for Mexico; and Bucheli and González (2007) for Uruguay.

An additional contribution of our analysis will be to distinguish between private and public sector workers within the formal segment of the labour force, but also by gender as the literature has shown significant differences in their reaction to local labour market conditions. The separate analysis of public sector workers is particularly interesting as their wages are fixed centrally and, typically, not through a bargaining process. Therefore, we expect their wages to be highly insensitive to local labour market conditions (Sanz-de-Galdeano and Turunen, 2006).

2. The Colombian Continuous Household Survey (CHS)

The data used in this paper come from the Continuous Household Survey (CHS). Our analysis focuses on the period 2002-2006,[4] a homogenous period which was characterized by a remarkable macroeconomic performance with very high GDP growth rates and controlled inflation.

This survey, which is carried out by the National Administrative Statistics Department, involves monthly household surveys with questions on labour force, unemployment, and other socioeconomic and sociodemographic characteristics of thirteen metropolitan areas in Colombia.[5] In fact, one of the reasons to use CHS microdata to carry out this study is that it provides information for the 13 functional areas[6] in Colombia, a territorial delimitation that is much more related to the concept of local labour markets than the usual administrative areas.[7]

We have used data from the second quarter of the year as it includes a special module on informality. Although there is not a consensus about how to define informality (and therefore how to measure it), the available information from the CHS permits to identify the most usual definition of informality (De Soto, 1987; Maloney, 2003) which is related to coverage by the

---

[4] The CHS was created to replace the National Household Survey (NHS), conducted since 1970 by the same institution. Later, in June 2006, the CHS underwent a major methodological change that gave rise to the Integrated Household Survey (GEIH). These changes make it difficult to extend the analysis prior to 2002 and after 2006.

[5] During the analyzed period, the population of those areas represented an average of 45.18% of the total population.

[6] Barranquilla, Bogota, Bucaramanga, Cali, Cartagena, Cucuta, Ibague, Manizales, Medellín, Monteria, Pasto, Pereira, Villavicencio.

[7] There are two main reasons for reporting our result by using these metropolitan areas: 1) these areas are the best way for communicating research results to policy-makers in Colombia, and 2) metropolitan areas are generally accepted in the literature as a good option for overcoming the Modifiable Areal Unit Problem (Gehlke and Biehl, 1934; Openshaw and Taylor, 1979).

social security system (formal workers) or not (informal workers).

There is also data on personal and job characteristics and labour income, among others. The availability of this broad individualised information, and the territorial detail makes it ideal for our purpose, as it is thus possible to analyse 13 territorial labour markets along the three months of the second quarter of the period 2002-2006. The panel dimension of the dataset also permits to control for unobservable regional characteristics, and, so, (at least, partially) avoid the omitted bias problem.

Moreover, an additional advantage of the CHS is that it was the official source for the analysis of regional labour markets, so unemployment rates at the territorial level can be directly calculated from this source.

Regarding labour income, we have combined information from monthly income and worked hours in order to obtain a measure of hourly wages that has been converted into real hourly wages using regional consumer price indices as deflators.[8]

Our final sample, after dropping individuals younger than 12 years old and older than 65, includes 174,908 worker and it is uniformly distributed along time. The share of informal workers in the sample is 41.0%, a figure that is quite close to the 38.4% estimated by OECD (2009). The 6.6% of workers are employed in the public sector, while the rest, 52.4% work in the formal private sector.

3. Econometric methodology

Our starting point to estimate the wage curve in Colombia is a Mincerian equation where the logarithm of individual income is regressed on a number of control variables related to personal and job characteristics and the local unemployment rate. A semi-logarithmic function, which, according to Mincer (1974), is the most appropriate functional form, is utilised. In particular, the logarithm of labour income is a function of a vector of individual and job characteristics and the local unemployment rate:

---

[8] The regional consumer price indices were obtained from the National Administrative Statistics Department (DANE). Besides the national level, the DANE reports the consumer price indices for the thirteen biggest cities in Colombia. Since each one of these cities is the core of a metropolitan area, we applied the consumer price index of the city to the whole metropolitan area.

$$w_{ijt} = f(z_{ijt}, u_{jt}) + e_{ijt} \qquad (1)$$

where w_{ijt} is the natural logarithm of the real hourly wage[9] of the individual i who lives in the metropolitan area j at time t, z_{ijt} is a set of individual factors that can affect wages of the individual, such as the level of schooling, potential experience, gender, occupation, activity sector, among others, u_{jt} is the unemployment rate in the metropolitan area j at time t and, finally, e_{ijt} is a random error term which is assumed to follow a normal distribution with mean zero and constant variance.

However, and before estimating equation (1), there is one potential problem that has to be taken into account: the possible omission of relevant variables at the territorial level. If relevant variables are not included, the coefficient associated with the unemployment rate (the only territorial variable in the regression) could pick up part of these effects when unemployment is correlated with these omitted variables. To consider this possibility, the usual approach is to include regional fixed effects, so equation (1) is augmented with regional dummy variables. Moreover, the panel dimension of our data allows us to include time fixed effects as additional control and additional time-varying regional variables. In particular, we have included regional productivity,[10] as there can be wage differences - not explained by previous control variables - related to the unequal efficiency levels among metropolitan areas or to the limited mobility of some factors. Moreover, this variable can capture the effects of different productive structures in each region, which probably are insufficiently controlled for by industry sector and occupational dummies. The relative advantages of panel data to cross-sectional analysis has been highlighted by Bratsberg and Turunen (1996) and, more recently, by Johnes (2007) in the context of the wage curve literature.

Taking all this into account, our final specification for the analysis is the following one:

$$w_{ijt} = f(z_{ijt}, u_{jt}, y_{jt}) + \delta_j + \gamma_t + e_{ijt} \qquad (2)$$

---

[9] This is the appropriate wage definition according to Card (1995) and Blanchard and Katz (1997).
[10] Data of Gross Domestic Product and employment were obtained from the National Administrative Statistics Department (DANE).

where the main difference with equation (1) is the inclusion of regional productivity (y_{jt}), regional fixed effects (δ_j) and time fixed effects (γ_t).

Having specified the model and defined the independent and explanatory variables, the next step consists of estimating equation (2). However, a difficulty arises because this equation includes an explanatory variable of interest (the unemployment rate of the metropolitan area j) that is defined at a higher level of aggregation (territory) than the dependent variable (individual). As Moulton (1986) shows, the estimation by ordinary least squares of this kind of equation will bias upward the values of the test of individual significance for this variable. In other words, the statistical significance of the unemployment rate, and thus the hypothesis of the presence of a wage curve may not be rejected due to the application of an inappropriate estimation procedure. To overcome this problem, we estimated equation (2) by grouping the data of the dependent variable and for each explanatory variable by calculating the average for individual groups (for the individuals in every territory j at time t). This procedure is known as 'cell-means' estimation. Therefore, our estimated equation is as follows:

$$\overline{w}_{jt} = f(\overline{z}_{jt}, u_{jt}, y_{jt}) + \delta_j + \gamma_t + \overline{e}_{jt} \tag{3}$$

where the notation of the variables is similar to that used in equation (2) and the subindex j and the subindex t are related to all the territories and time periods considered, respectively. However, when working with grouped data, the OLS estimator is unbiased but inefficient (Greene, 1998, 374-6) and, for this reason, estimated standard errors will be adjusted for the presence of heteroskedasticity. The results of estimating equation (3) using CHS microdata and disaggregating between different workers groups are shown in the next section.

4. Empirical evidence

We first estimated by Ordinary Least Squares a Mincer equation for the full sample without including the local unemployment rate. The most relevant results are shown in Table A.1 of the annex.[11] The results indicate that the wages of women are substantially lower than those of men, and that workers in the informal sector earn a 30% less than workers with similar characteristics in the formal private sector. On the opposite, workers in the public sector earn

---

[11] Full results are available from the authors on request.

nearly a 30% more. Variables related to human capital – schooling years and years of potential experience - were significant and showed the expected signs, detecting a positive relationship between schooling and wages. More educated workers receive higher wages. The estimated returns to schooling are around the 7% and slightly higher for men than for women. The accumulation of professional experience also had a positive effect on wages, although the concavity of the relationship- revealed by the negative value of the coefficient of the variable square of the experience - indicated decreasing returns for investment in specific human capital and even the existence of a starting age from which additional experience had a negative influence on wages. The regression also includes additional controls for activity sectors, occupations and regional and time fixed effects. The results, which are available from the authors on request, are similar to the ones obtained in this kind of empirical exercise.

Once the economic and statistical significance of the OLS estimates have been assessed, we have proceed to the "cell means" estimation of the Mincer equation with the regional unemployment rate and regional productivity as additional explanatory variables. The results for the wage curve coefficient, the elasticity of individual wages to local unemployment, are shown in table 1.

The first column in table 1 shows the results for wage curves without separating men from women. The results for all workers are shown in the first row, while the results for informal and formal workers are shown in the following rows. A significant and negative relationship between individual wages and the contemporaneous regional unemployment rate is observed, confirming the results obtained on wage curves for other countries. Moreover, the value of the coefficient, which can be interpreted as the elasticity of the curve, is –0.07, a value close to the –0.10 found by Blanchflower and Oswald (1994). However, when focusing on informal and formal workers, the results suggest that only wages of informal workers react to local labour market conditions. Their elasticity to regional unemployment is significantly higher than the usual -0.10: -0.18, while the elasticity for workers in the formal sector is not significantly different from zero for both public and private sector workers. The aggregate result seems to be related to the higher response of informal workers than to an average effect of the different groups.

The results of estimating disaggregated wage curves by gender are shown in the second and third columns for men and women, respectively. A more robust wage curve is found for men

7

with a significant elasticity of –0.07 as for women, no effect of unemployment on wages is found.[12] However, the results for informal workers are different as in both cases, a wage curve is found. One potential explanation of this result is that unemployment not only affects wages but also participation decisions: a high level of unemployment increases the number of discouraged workers, thereby reducing the labour supply and increasing wages. Where this effect is low –for example, among men- the initial negative effect on wages will clearly dominate, but if this effect is relevant –for example, for women- both effects will be opposite showing no evidence of a wage curve.

Table 1. Cell means estimation of the wage curve for different groups of workers

		All workers	Male	Female
All workers		-0.0707**	-0.0761**	-0.0600
		[0.0276]	[0.0268]	[0.0369]
Informal workers		-0.1790**	-0.1810*	-0.2000*
		[0.0799]	[0.0839]	[0.109]
Formal workers	Private sector	-0.0598	-0.0941*	-0.0157
		[0.0343]	[0.0433]	[0.0555]
	Public sector	0.0145	0.1670	-0.0600
		[0.0658]	[0.104]	[0.109]

All models include controls for activity sector, occupation, regional productivity, regional and time fixed effects, and gender, informal sector and public sector, when possible.

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Summarising, our results have shown the existence of wage curve for workers in the informal sector (both men and women) and for men working in the private formal sectors, although the value of the elasticity of wages to unemployment is significantly lower than for the rest of workers. No evidence of a wage curve is found for public sector workers.

---

[12] A result which is in line with the literature: see Janssens and Konings (1998).

5. Final remarks

In this paper, first, we have estimated a wage curve for the Colombian economy using data for the recent economic expansion, and second, we have estimated disaggregated wage curves between informal and formal workers in order to test if the response of wages to labour market conditions is higher for the first, a controversial issue in the scarce literature on the topic for developing economies.

The obtained results show the existence of a wage curve with a negative slope for the Colombian economy. Using information on metropolitan areas, the estimates of the elasticity of individual wages to local unemployment rates was -0.07 a value that is very close to those obtained for other countries. Therefore, this result seems to confirm that differences in the institutional framework among countries does not seem to affect (or only slightly affects) the sensitivity of individual wages to local labour market conditions.

However, the disaggregation of statistical information for formal and informal workers has shown significant differences among both groups of workers. In particular, for the less protected groups of the labour market, informal workers (both men and women), a high negatively sloped wage curve was found. This result is consistent with the conclusions from efficiency wage theoretical models and should be taken into account when analysing the functioning of regional labour markets in developing countries.

6. References

Babecký, J.; Ramos, R.; Sanromá, E. (2008), "Meta-analysis on microeconomic wage flexibility (wage curve)", Sozialer Fortschritt, vol. 57 (10-11), pp. 273-279.

Barros, R. P.; Mendonca, R. (1996). Flexibilidade do mercado de trabalho brasileiro: uma avaliação empírica, in Camargo, J. M. (ed.), Flexibilidade do mercado de trabalho no Brasil. Rio de Janeiro: FGV Editora, pp. 157-201.

Berg, J. and Contreras, D. (2004). Political-Economic Regime and the Wage Curve: Evidence from Chile, 1957-1996. *International Review of Applied Economics*, 18,(2), pp. 151-165.

Blanchard, O. and Katz, L.F. (1997), What We Know and Do Not Know about the Natural Rate of Unemployment, *Journal of Economic Perspectives*, 11(1), pp. 51-72.

Blanchflower, D.G. and Oswald, A.J. (1990). The Wage Curve. *Scandinavian Journal of Economics*, 92(2), pp. 215-235.

Blanchflower, D.G. and Oswald, A.J. (1994), The wage curve, MIT Press, Cambridge, MA.

Blanchflower, D.G. and Oswald, A.J. (2005). The wage curve reloaded. National Bureau of Economic Research Working Paper No.11338, Cambridge, MA.

Bratsberg, B. and Turunen, J. (1996), Wage Curve Evidence from Panel Data, *Economics Letters*, 51, pp. 345-353.

Bucheli, M. and González, C. (2007). An Estimation of the Wage Curve for Uruguay. Working Paper 11/07. Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República, Uruguay.

Card. D. (1995). The Wage Curve: A Review, *Journal of Economic Literature*, 33(2), pp. 285-299.

Castro, D. (2006). Curva salarial: una aplicación para el caso de México, 1993-2002. *Estudios Económicos*, 21(2), pp. 233-273.

De Soto, H. (1987). El Otro Sendero. La Revolución Informal, Instituto Libertad y Democracia, Lima.

Galiani, S. (1999). Wage determination in Argentina: An econometric analysis with methodology discussion. Wolfson College, Oxford University.

Gehlke, C. and Biehl, K. (1934). Certain effects of grouping on upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29:169-170.

Greene, W.H. (1998), Econometric Analysis. 3rd Edition, Prentice Hall, Upper Saddle River.

Janssens, S. and Konings, J. (1998), One more wage curve: the case of Belgium, *Economics Letters,* 60, pp. 223-227.

Johnes, G. (2007), The Wage Curve Revisited: Estimates from UK Panel, *Economics Letters*, 94, pp. 414-420.

Maloney, W. (2003). Informality Revisited. *World Bank, Policy Research Working Paper* 2965.

Mincer, J. (1974). Schooling, Experience and Earnings. Columbia University Press, New York.

Moulton, B.R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics*, 32, pp. 385-397.

Nijkamp, P. and Poot, J. (2005). The last word on the wage curve? A meta-analytic assessment. *Journal of Economic Surveys*, 19, pp. 421-450.

OECD (2009), Is Informal Normal? Towards More and Better Jobs in Developing Countries, OECD Development Center.

Openshaw, S. and Taylor, P. (1979), A million or so correlation coefficients, In N. Wrigley (Ed.), Statistical Methods in the Spatial Sciences, London, Pion, pp. 127–144.

Sánchez, F. and Nuñez, J. (1998). La curva de salarios para Colombia. Estimaciones de las relaciones entre desempleo, la inflación y los ingresos laborales, 1994-1996. DNP. planeación y Desarrollo. Vol. XXIX Julio-Septiembre.

Sanz-de-Galdeano, A. and Turunen, J. (2006), The Euro Area Wage Curve, *Economics Letters*, 92, pp. 93-98.

7. Annex

Table A.1. OLS estimates of the Mincer equation

Log of real hourly wage	All workers	Male	Female
Male	0.102***		
	[0.00385]		
Married	0.0319***	0.0367***	0.0219***
	[0.00345]	[0.00574]	[0.00544]
Children	-0.0178***	-0.0257***	-0.00673
	[0.00593]	[0.00723]	[0.0101]
Main income provider	0.136***	0.123***	0.126***
	[0.00364]	[0.00557]	[0.00631]
Schooling years	0.0698***	0.0706***	0.0681***
	[0.000600]	[0.000755]	[0.000984]
Potential experience	0.0200***	0.0236***	0.0166***
	[0.000467]	[0.000622]	[0.000712]
Potential experience^2	-0.000260***	-0.000313***	-0.000213***
	[9.16e-06]	[1.18e-05]	[1.45e-05]
Informal	-0.324***	-0.336***	-0.316***
	[0.00394]	[0.00499]	[0.00636]
Public sector	0.296***	0.238***	0.384***
	[0.00629]	[0.00950]	[0.00875]
Activity sector dummies	Yes	Yes	Yes
Occupational dummies	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Observations	174908	96156	78752
R-squared	0.445	0.463	0.431

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1