Microeconometric Decompositions of Aggregate Variables

An application to Labor Informality in Argentina

Leonardo C. Gasparini

UNLP y FIEL

Abstract

This paper illustrates the use of microeconometric decomposition techniques to characterize changes in aggregate variables. In particular, I study the effect of changes in the employment structure on the labor informality rate for salaried workers in the Greater Buenos Aires area. To that aim I compute the difference between the informality rate at moment *t´* and the rate that results from combining the population at moment *t* with the parameters estimated at moment *t´* that link observable individual characteristics to the informality decision. The paper concludes that the deep change of the employment structure in Argentina during the 1980s and the 1990s has had a significant but minor effect on the labor informality rate.

Keywords: decomposition, informality, employment, Argentina

JEL classification: J2, J3, H2, H3

1. Introduction

Changes in an aggregate variable could be the consequence of changes in individual characteristics and/or in the way these characteristics are linked to economic behavior. Take the case of the unemployment rate, which can be expressed as a weighted average of the unemployment rates of different sociodemographic groups with different propensities to be unemployed. Changes in the unemployment rate could be due to changes in the share of each group in the population and/or changes in the way sociodemographic characteristics are linked to the probability of being unemployed. To separate out both components, microeconometric decomposition techniques could be applied. It is a simple idea. Let Y be an aggregate variable obtained as an average of an individual variable Y_i. Suppose that variable Y_i is a function of individual characteristics X_i, β being the corresponding parameters vector. The change in the value of the aggregate variable Y is then determined by changes in individual characteristics X_i and by changes in vector β. The basic idea of the microeconometric decomposition technique consists in simulating the value of Y_i at moment t', keeping alternatively fixed the individual characteristics and the parameters at moment t. If, for example, the value of Y simulated according to the individual characteristics at moment t and the parameters at t' is not significantly different from the real value of Y at t', it can be concluded that changes in individual characteristics are not really important to explain changes in variable Y.

Microeconometric decomposition techniques have been mainly applied to the study of discrimination (Blinder, 1973; Oaxaca, 1973; Oaxaca and Ramson, 1994) and inequality (Juhn, Murphy and Pierce, 1993; Blau and Kahn, 1996; Bourguignon, Ferreira and Lustig, 1998). This paper combines both strands by focusing on decompositions on means like the discrimination literature, but going beyond the linear regression model originally used by Blinder (1973) and Oaxaca (1973), as in the literature initiated by Bourguignon *et al*. (1998).

The article applies the technique sketched above to the study of a particular problem: the dramatic increase in the rate of informal employment in Argentina during the 1980s and 1990s. The percentage of salaried informal workers in Greater Buenos Aires was 19.9% in 1980. This percentage rose to 29.3% in 1990 and reached a record of 38.4% in 1998. (Gasparini, 2000). This pattern is similar in the rest of the urban labor markets of the country. The following argument has gained popularity to understand this phenomenon. Argentina has experienced a deep transformation process in its employment structure during the last decades: the fraction of people working in small firms has increased, the seniority in jobs has been reduced, the percentage of part-time jobs has increased, the industry has lost importance at the expense of most services, and the participation of different sociodemographic groups in total employment has significantly changed. In this context the increase in the aggregate rate of labor informality may be the result of increases in the share of occupations or activities with a high propensity to informal arrangements. The higher rate of informality could simply be the consequence of an employment structure biased towards positions, activities or sociodemographic groups with a higher propensity to informality.

This article intends to evaluate the empirical validity of this argument using microeconometric decomposition techniques. The aggregate informality rates required to implement the methodology derive from estimations of the labor formal-informal individual choice based on a model of equalizing differences for job amenities (see Rosen, 1986).

The results of this article suggest that, even though the deep change in the employment structure in Argentina is a significant factor, its role in the explanation of the labor informality rate is minor. The regional analysis arrives at the same conclusion. The large differences in the productive and employment structure among regions give account of only a small fraction of the substantial differences in the informality rates.

The rest of the article is organized in the following way. Section 2 describes the methodology to characterize changes in aggregate variables. In section 3 an application to the case of informality in Greater Buenos Aires is suggested, in section 4 a formal-informal choice model is estimated, and in section 5 the results of the decompositions are shown. Section 6 concludes with some final comments.

2. The methodology

Let Y_t be an aggregate variable at moment *t* defined as the average of individual variables Y_{it}

(2.1) $$Y_t = \sum_{i=1}^{N} \frac{Y_{it}}{N}$$

Suppose that the population can be classified into *J* groups formed by individuals who share the same value of all the observable variables that potentially affect *Y*. In this case,

(2.2) $$Y_t = \sum_{j=1}^{J} \theta_{jt} Y_{jt}$$

where Y_{jt} is the average of *Y* in group *j* and θ_{jt} indicates the share of this group in the population at moment *t*. The change in the aggregate variable between *t* and *t'* can be written as:

(2.3) $$\Delta Y \equiv Y_{t'} - Y_t = \sum_{j=1}^{J} \theta_{jt} \Delta Y_j + \sum_{j=1}^{J} Y_{jt'} \Delta\theta_j$$

The first term captures the effect of the changes in *Y* in each group, while the second term seizes the changes in the shares of each group. Although theoretically the implementation of (2.3) is simple, in practice the number of observationally equal groups *J* may be too large - it can even coincide with *N* -, and hence turn the decomposition senseless. Additionally, the groups present at *t* will probably not exist at *t´*, even with panel data, as the characteristics of individuals change over time. These problems are particularly probable if there are many relevant characteristics, and if each one has many categories (or worse, if they are continuous). The usual practice is to divide the population into groups defined in terms of only some characteristics, which in turn are classified into few categories. The result is a distortion in the decomposition (2.3) as differences within each of the groups are ignored.

This article uses a decomposition technique based on micro data with an econometric modeling of the individual decisions to characterize the change in an aggregate variable. This microeconometric technique can soothe the problems of the aggregate decompositions. The method begins by modeling Y_{it} as a function *H* of a vector of observable individual characteristics X_{it}, a parameter vector β_t and unobservable factors ε_{it}. Assuming that the form of function *H* does not change over time, Y_{it} can be written as:

(2.4) $$Y_{it} = H(X_{it}, \beta_t, \varepsilon_{it})$$

The function *H* (.), the parameters β_t and the random terms are unobservable. The usual procedure of estimation implies assuming some functional form for *H* (.), estimating the parameters β_t through some econometric technique and obtaining the unobservable factors (and their returns) as residuals. Formally, the individual value of Y_{it} can be expressed as:

(2.5) $$Y_{it} = \hat{H}\left(X_{it}, \hat{\beta}_t\right) + \hat{\varepsilon}_{it}$$

where $\wedge$ stands for estimated (or simulated in the case of the function *H*). From the preceding definitions and assuming for simplicity $N_t = N_{t'} = N$, the change in the value of *Y* between *t* and *t´* is

(2.6) $$\Delta Y = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it'}, \hat{\beta}_{t'}\right) + \hat{\varepsilon}_{it'}\right) - \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it}, \hat{\beta}_{t}\right) + \hat{\varepsilon}_{it}\right)$$

The aggregate variable can change basically due to changes in the observable characteristics of the population and to changes in the β parameters.

Characteristics effect

What would have been the change in *Y* if only the observable characteristics of the population had changed between *t* and *t´*? The "characteristics effect" measures this counterfactual change. Taking *t* as the base year this effect is computed as the difference between the value that would have resulted if only the observable characteristics of the population had changed between *t* and *t´* and the value of the aggregate variable at *t*. Analytically,

(2.7) $$CE_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it'}, \hat{\beta}_{t}\right) + \hat{\varepsilon}_{it}\right) - \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it}, \hat{\beta}_{t}\right) + \hat{\varepsilon}_{it}\right)$$

Equation (2.7) isolates the effect of changing individual characteristics from *t* to *t'*, keeping the rest constant at time *t* values. The same exercise can be done taking *t'* as the base period. In that case, the characteristics effect (denoted as CE_2) is computed as the difference between the observed aggregate variable in *t'* minus the simulated value with time *t* characteristics and time *t'* parameters and residuals.

Parameters effect

This effect captures the change in *Y* as a consequence of changes in parameters β, keeping all the rest fixed. Taking *t* as the base year it can be written as,

(2.8) $$PE_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it}, \hat{\beta}_{t'}\right) + \hat{\varepsilon}_{it}\right) - \frac{1}{N}\sum_{i=1}^{N}\left(\hat{H}\left(X_{it}, \hat{\beta}_{t}\right) + \hat{\varepsilon}_{it}\right)$$

In a similar way, this effect can be computed by taking *t'* as the base year (denoted as PE_2). It is easy to show that the change in the aggregate variable equals the sum of the averages (changing the base year) of the characteristics and parameters effects plus a residual that is likely to be insignificant.

(2.9) $$\Delta Y = \frac{(CE_1 + CE_2)}{2} + \frac{(PE_1 + PE_2)}{2} + R$$

where $R = \frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_{it'} - \frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_{it}$. By construction, *R* is zero in standard regression models estimated by OLS, which include a constant. But even in models where *R* could be different from zero (e.g., probit-MLE, median regression, etc.), in most practical cases it will be insignificant.

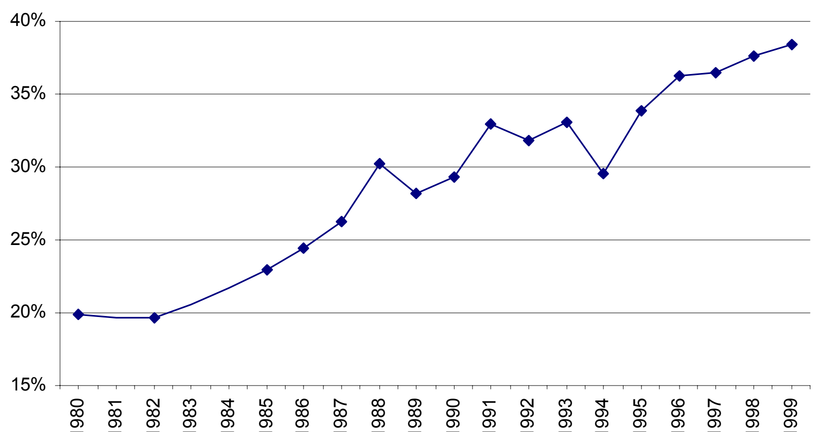
The only input required by the decomposition is a matrix of observed and estimated aggregate variables. In the following sections, this methodology is illustrated with an application to a particular case: the increase of labor informality in Argentina. The methodology can be easily extended to the study of other aggregate variables, such as the rate of unemployment, the average wage, the rate of medical coverage or the rate of school attendance.

3. Informality in Greater Buenos Aires

There are two alternative views of the concept of informality. The first one refers to marginal jobs of low productivity and to subsistence economic units with low or null accumulation capacity. The alternative view emphasizes the illegal or non-regulated character of the employment relationship. In this article, the second notion of informality is adopted. Specifically, an informal worker is defined as a person who declares not to have retirement benefits in its main job.[i]

Figure 3.1 shows the main phenomenon to be studied in this article. According to information from the Permanent Household Survey (EPH) of Argentina, the informality rate of salaried workers in Greater Buenos Aires (GBA) grew rapidly and steadily during the eighties, it stabilized in the first half of the nineties and returned to its ascending pattern from 1995 on.[ii] The percentage of informal salaried workers was around 20% in the early eighties. This figure grew to 30% in the late eighties and to nearly 40% in the late nineties.[iii]

Figure 3.1
Informality rate for salaried workers
Greater Buenos Aires, 1980-1999



Source: own calculations based on the EPH.

The Greater Buenos Aires labor market has experienced considerable changes in the last decades. Table 3.1 shows the wage-earners employment structure in four selected years: 1980, 1986, 1992 and 1998. The sociodemographic employment structure has changed in favor of women, people older than 60 and workers with higher education. The productive structure, and therefore that of employment, has varied in favor of most services and against the industry. Table 3.1 shows a trend towards employment in smaller establishments.[iv] The seniority in employment has been reduced and part-time employment has increased, especially in the nineties.

Changes in the employment structure can affect the rate of labor informality if the propensity of informal labor arrangements is different among different types of employment and worker characteristics. Table 3.2 documents the substantial differences in informality rates among groups. In particular, labor informality is relatively high for women and low seniority workers and in part-time and small firms jobs, whose relative shares have increased in the last two decades.

Table 3.2 gives rise to an alternative explanation of the growth in the informality rate by showing an increase in the share of informal workers in practically all the sociodemographic, sectoral, and labor groups considered. Even with a fixed employment structure, these changes naturally translate into a higher aggregate labor informality rate.

Summarizing, labor informality has strongly increased in the Greater Buenos Aires area during the eighties and nineties. This phenomenon appears to be the consequence of (i) changes in the employment structure towards jobs with a higher propensity to informal arrangements and (ii) generalized increases in informality within each group.

The decomposition methodology detailed in section 2 helps us to evaluate the relative importance of these two channels. The methodology requires the modeling of individual choices for labor informality.

4. A formal-informal choice model

In this section the labor informality choice is modeled and the empirical estimation strategy is presented.[v] Suppose a salaried worker *i* with a separable utility function

(4.1) $$U_i(x_i, B_i) = u(x_i) + t_i B_i$$

where x_i denotes the consumption of a numeraire good and B_i is a binary variable that indicates whether or not the worker receives social security. The social protection is understood in a broad sense, including retirement payments, health insurance, holidays, severance payments, labor insurance, family allowances, and other benefits. For simplicity, it is assumed that benefits are the same for every worker but preferences for these benefits, represented by parameter t_i, vary across individuals. The individual spends his/her labor wage W_i on the numeraire good and social security payments. The individual labor supply is elastic at the reservation wage W_i^r.

Firms in this economy provide social benefits, which implies a wedge between the wage and the labor cost. The provision of social protection does not only generate direct costs but also implies a loss of flexibility in labor contracts, which implies lower profits. The cost of providing benefits to a worker is labeled as C_i and may vary among workers. In a non-regulated labor market firms can entirely shift this burden, providing social benefits only to those workers who are willing to pay for them through a wage reduction. Instead, if the government requires compulsory benefits provision and the firms cannot elude the regulation, the shifting will be partial. There are workers who prefer not to have any social protection at cost C_i. Consequently, the firm has to bear the difference between the cost of the benefits C_i and the worker's willingness to pay for them D_i.[vi]

The firms have a way to try to escape from the mandatory provision of social protection: labor informality. The cost of this policy comes from a positive probability *p* of being caught and fined for having irregular labor agreements.[vii] In this case, the firm does not only have to pay C_i, a part of which –the worker's willingness to pay- can be recovered, but it must also pay a fine *M*. When facing the decision of hiring another worker with $D_i \leq C_i$, the firm must compare the net profits of hiring the new worker as a formal one

(4.2) $$VMP_i(L) - W_i^r - C_i + D_i$$

(where $VMP_i(L)$ is the value of marginal product when L workers are already hired and i is the marginal worker), with the net profits of hiring the new worker as an informal one

(4.3) $$VMP_i(L) - W_i^r - p.(C_i - D_i + M)$$

From (4.2) and (4.3) a risk-neutral firm will offer social benefits to its worker i if

(4.4) $$D_i \geq C_i - \frac{p}{1-p}.M$$

Hence, for an individual to receive and accept a work offer with social benefits, it must be the case that

(4.5) $$t_i \geq u(W_i^r) - u(W_i^r - C_i + (p/(1-p)).M)$$

Except in the case of an expected fine $p.M$ equal to zero, the proportion of covered individuals will be greater than the one that would emerge from the case of voluntary protection. However, except when $p.M$ is very large, this proportion will be less than 1, the result of a compulsory protection without evasion.[viii]

For empirical implementation the following stochastic structure can be assumed:

(4.6) $$t_i = \beta' X_i + \mu_i$$

(4.7) $$C_i = \delta' Z_i + \nu_i$$

where X_i and Z_i are vectors of explanatory variables, β and δ are vectors of parameters and μ_i and ν_i are stochastic terms. A first order approximation to the RHS of (4.5) is

(4.8) $$\gamma_0 + \gamma_1 W_i^r + \gamma_2 C_i + \gamma_3 p + \gamma_4 M$$

Replacing (4.6), (4.7) and (4.8) in (4.5) yields

(4.9) $$\varepsilon_i = \mu_i - \gamma_2 \nu_i \geq \gamma_0 + \gamma_1 W_i^r + \gamma_2 \delta' Z_i + \gamma_3 p + \gamma_4 M - \beta' X_i$$

If the random term ε_i is higher than the RHS of (4.9), the worker receives social benefits. Assuming that μ_i and ν_i are normally distributed, we have a *probit* model, which can be estimated by maximum likelihood methods. This is a binary choice model where the dependent variable equals 1 if the salaried worker receives social benefits from the firm and 0 otherwise.

The preceding model ignores the family structure within which the social coverage decision is usually made. In particular, it is possible that the income effect, which in this model stems from the individual wage, is better captured by household income. For this reason and also because of the lack of information regarding the individual reservation wage, in the estimations the equivalized household income is taken as the relevant income variable, instead of W_i^r. Gender, marital status, age (and its square) and education are included as individual sociodemographic characteristics that can affect the demand for social benefits. Additionally, the size of the firm, job seniority, a dummy for part-time jobs and the sector of activity are included to capture both the factors that modify the cost of protection provision and the probability of being fined.

The results of the estimation are displayed in Table 4.1. Figures 4.1 to 4.4 take these results to show the conditional probabilities of labor informality as a function of age, educational level, size of the firm and job seniority. The curve that relates informality with age is U shaped, indicating higher probabilities of irregular jobs for younger and older workers (Figure 4.1). There are no clear trends between years, apart from a sharp increase in the constant. Figures 4.2 to 4.4 reproduce the result of the increase in the constant over time. Controlling for other factors, the probability of informality has increased more for less educated workers (Figure 4.2), smaller firms (Figure 4.3) and for workers with less seniority (Figure 4.4). However, these differential changes seem to be minor when compared to the vertical shift of the curves.

5. Results

To implement the decomposition of section 2 a matrix of observed and simulated informality rates is required. Table 5.1 shows that matrix for the Greater Buenos Aires. In order to compute it, the third probit model of Table 4.1 is used. According to Table 5.1 in 1980 the labor informality rate was 18.9%.[ix] That rate would have been 22.4% if the parameters that link individual characteristics to the labor informality decision had been those of 1986. With fixed parameters at their values in 1980 and the population of 1986 the rate would have been 19.0%.

Table 5.1
Simulated informality rates
Greater Buenos Aires, 1980, 1986, 1992 y 1998

	Parameters of year			
	1980	1986	1992	1998
1980	0.189	0.224	0.289	0.330
1986	0.190	0.218	0.285	0.323
1992	0.221	0.253	0.315	0.358
1998	0.230	0.267	0.325	0.371

Source: own calculations based on the EPH. Each informality rate is obtained using the parameters of model 3 in Table 4.1.

Reading the matrix 5.1 by column allows a first evaluation of the characteristics effect. Keeping the parameters fixed in a given year, the individual and employment characteristics of the population have changed so as to generate increasing informality rates. This conclusion applies to each column. However, the changes generated by the differences in characteristics do not seem to be large. A different conclusion is reached when examining Table 5.1 by row, which gives an idea of the parameters effect. In all the cases, the simulated labor informality rate significantly grows between periods.

The results of the decompositions are shown in Table 5.2.[x] In columns (i) and (ii) the initial year is taken as base year, in columns (iii) and (iv) the final year is the base year, and in columns (v) and (vi) the average of the two simulations is reported. Column (vii) shows the actual change in the rate of labor informality in each period. That rate, for instance, increased 9.6 percentage points between 1986 and 1992. The average of computing equation (2.7) changing the base year gives a value of 3.2%, while the average of computing equation (2.8) changing the base year is 6.5%.[xi]

Table 5.2
Decomposition of the informality rate
Greater Buenos Aires, 1980, 1986, 1992 y 1998

	Initial year as base		Final year as base		Average		Actual change
	Characteristics	Parameters	Characteristics	Parameters	Characteristics	Parameters	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
80-86	0.001	0.035	-0.006	0.028	-0.002	0.032	0.029
86-92	0.034	0.067	0.029	0.062	0.032	0.065	0.096
92-98	0.010	0.043	0.013	0.046	0.011	0.044	0.056
80-98	0.041	0.140	0.041	0.140	0.041	0.140	0.182

Source: own calculations based on the EPH. See text for a description of the methodology.

From Table 5.2 it can be seen that the characteristics effect has always been lower than the parameters effect. Additionally, it is interesting to see that, except for one case, all the effects in all the periods are positive, that is, they contribute to the increase in labor informality. Between 1980 and 1986 the change in the parameters explains the whole increase in the informality rate. In fact, the characteristics effect is negative on average. The transformation of the productive and employment structure in Argentina between 1986 and 1992 translated into a stronger characteristics effect. However that effect does not reach half the value of the parameters effect. The relative importance of the changes in the employment structure on the labor informality rate is smaller in the 1992-1998 period.

Summarizing, although the change in the employment structure had an increasing effect on the labor informality rate, its quantitative importance seems secondary compared to the change in the parameters of the formal-informal individual choice. A significant increase in the propensity to evade labor taxes seems to be the main force behind the increase in the rate of labor informality.

Characterization of regional differences

There is a great variability in the rates of labor informality among the urban labor markets covered by the EPH. Table 5.3 shows the percentage of informal salaried workers in October, 1997 in many Argentine cities.

Table 5.3
Informality rates for salaried workers
Argentine cities, 1997

Cities	informality rate
Comodoro Rivadavia	0.27
Gran Buenos Aires	0.37
Jujuy	0.43
La Pampa	0.31
La Plata	0.35
Neuquén	0.33
Paraná	0.33
Río Gallegos	0.21
Salta	0.43
San Juan	0.38
San Luis	0.32
Santiago del Estero	0.46
Tierra del Fuego	0.21

Source: own calculations based on the EPH.

The decomposition methodology provides a tool to evaluate the relevance of the different sociodemographic, productive and employment structure among Argentine cities as a source of differences in labor informality. To simplify, three regions are considered: the South, the Greater Buenos Aires and the Northwest (NOA).[xii]

In Table 5.4 the informality rates in one region are simulated using the parameters of a probit model of the formal-informal choice of another region. The rate of labor informality in the South is 22.8%. With the estimated parameters of NOA, that rate would have been much higher:

37.5%. This suggests again a strong parameters effect and a less important characteristics effect.

Table 5.4
Simulated rates of informality
South, GBA and NOA regions, 1997

	Parameters of region		
	South	GBA	NOA
South	0.228	0.320	0.375
GBA	0.285	0.361	0.445
NOA	0.301	0.363	0.417

Source: own calculations based on the EPH.

These presumptions are confirmed in Table 5.5. The differences in the sociodemographic and employment structure of the different regions explain a significant, though minor, part of the differences in the rates of informality.

Table 5.5
Decompositions of the rate of informality
South, GBA and NOA regions, 1997

	Initial region as base		Final region as base		Average		Actual difference
	Characteristics	Parameters	Characteristics	Parameters	Characteristics	Parameters	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
South-GBA	0.057	0.092	0.041	0.076	0.049	0.084	0.133
South-NOA	0.073	0.147	0.042	0.116	0.058	0.132	0.189
GBA-NOA	0.002	0.084	-0.028	0.055	-0.013	0.069	0.056

Source: own calculations based on the EPH. See text for a description of the methodology.

6. Concluding remarks

This article shows the usefulness of microeconometric decompositions to examine the reasons behind the changes in aggregate variables. These decompositions allow to separate out the effect of changes in individual characteristics from the effect of changes in the parameters that link these characteristics to economic behavior. The methodology is used to evaluate the hypothesis that changes in the employment structure explain much of the increase in the rate of labor informality in Argentina during the 1980s and 1990s. The study concludes that the deep transformation in the employment structure in Greater Buenos Aires has had a significant, but minor, positive effect on the rate of informality. According to the available evidence, this rate has grown to a great extent due to the increase in the propensity to evade labor taxes in practically all the sociodemographic and labor categories considered.

Acknowledgements

This article started as part of a project on the Hidden Economy made by Fundación de Investigaciones Económicas Latinoamericanas (FIEL). I am grateful for the excellent research assistance of Diego Margot, the observations made by participants at Seminars in FIEL and UNLP, and by Juan Luis Bour, Marcela Cristini, Walter Sosa Escudero and Nuria Susmel. I also

thank Florencia Jaureguiberry, who helped me with the translation from Spanish. Mistakes are my own responsibility.

References

Arango, C. and W. Maloney (2000). Employment, Informality and Poverty in the Argentine Labor Market. In The World Bank Report 19992-AR, *Argentina: Poor People in a Rich Country*, Background Paper 3.

Blau, F. and L. Kahn (1996). International Differences in Male Wage Inequality: Institutions versus Market Forces. *Journal of Political Economy* 104 (5), pp. 791-837.

Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimate. *The Journal of Human Resources* VIII (4), pp. 436-453.

Bourguignon, F., Ferreira, F. and N. Lustig, N. (1998). The Microeconomics of Income Distribution Dynamics in East Asia and Latin America. Unpublished IDB-World Bank Research Proposal.

Carpio, J., Klein, E. and I. Novacovsky (eds.) (2000). *Informalidad y Exclusión Social*. Fondo de Cultura Económica de Argentina, Buenos Aires.

de la Rica, S. and T. Lemieux (1994). Does Public Health Insurance Reduce Labor Market Flexibility or Encourage the Underground Economy? Evidence from Spain and the United States. In Blank, R. (ed.) *Social Protection versus Economic Flexibility: Is There a Trade-Off?*, NBER Comparative Labor Market Series, University of Chicago Press, Chicago and London, pp. 265-299.

Gasparini, L. (2000). La Informalidad Laboral en la Argentina. Evolución y Caracterización. In FIEL, *La Economía Oculta en la Argentina*, Buenos Aires, pp. 161-193.

Juhn, C, Murphy, K. and Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy* 101 (3), pp. 410-442.

Maloney, W. (1999). Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico. *The World Bank Economic Review* 13 (2), pp. 275-302.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Market. *International Economic Review* 14 (3), pp. 693-709.

Oaxaca, R and M. Ramson (1994). Discrimination and Wage Decomposition. *Journal of Econometrics* 61 (1), pp. 5-21.

Rosen, S. (1986). The Theory of Equalizing Differences. In Ashenfelter, O. and R. Layard (eds.), *Handbok of Labor Economics*, vol. 1, Elsevier Science B.V-North Holland, pp. 641-692 .

Table 3.1
Distribution of salaried workers
Greater Buenos Aires, 1980, 1986, 1992 and 1998

	1980	1986	1992	1998
Gender				
Men	0.66	0.62	0.62	0.60
Women	0.34	0.38	0.38	0.40
Age				
0-17	0.06	0.04	0.04	0.02
18-24	0.20	0.17	0.21	0.20
25-39	0.36	0.39	0.36	0.38
40-59	0.34	0.35	0.34	0.34
60-	0.03	0.05	0.05	0.06
Marital status				
Single	0.42	0.39	0.41	0.44
Married	0.58	0.61	0.59	0.56
Education				
Primary school drop-outs	0.18	0.14	0.09	0.07
Primary school graduates	0.36	0.32	0.31	0.26
High school drop-outs	0.19	0.20	0.21	0.21
High school graduates	0.14	0.15	0.18	0.18
College drop-outs	0.07	0.09	0.10	0.14
College graduates	0.06	0.09	0.11	0.14
Sector				
Low tech industry	0.12	0.12	0.11	0.07
Medium and high tech industry	0.24	0.20	0.16	0.13
Transport and utilities	0.08	0.07	0.08	0.10
Construction	0.08	0.04	0.04	0.06
Retail and wholesale trade	0.14	0.14	0.17	0.16
Business and financial services	0.07	0.09	0.09	0.12
Public Administration	0.05	0.06	0.06	0.07
Social services	0.09	0.11	0.13	0.15
Other services	0.12	0.17	0.14	0.14
Size of the firm (employees)				
1	0.08	0.09	0.10	0.09
From 2 to 5	0.17	0.18	0.24	0.22
From 6 to 25	0.25	0.24	0.25	0.26
From 26 to 100	0.21	0.21	0.20	0.20
From 101 to 500	0.15	0.14	0.12	0.14
More than 500	0.14	0.14	0.09	0.09
Seniority				
Until 3 months	0.11	0.10	0.12	0.15
From 3 to 6 months	0.07	0.06	0.06	0.07
From 6 months to 1 year	0.13	0.12	0.14	0.13
From 1 year to 5 years	0.33	0.30	0.29	0.32
From 5 years to 10 years	0.15	0.18	0.16	0.15
From 10 years to 20 years	0.11	0.14	0.15	0.13
More than 20 years	0.10	0.10	0.08	0.06
Weekly hours of work				
From 1 to 19	0.03	0.04	0.05	0.09
From 20 to 29	0.06	0.07	0.08	0.09
From 30 to 40	0.35	0.29	0.35	0.28
From 41 to 45	0.18	0.20	0.15	0.14
From 46 to 61	0.27	0.29	0.27	0.27
More than 62	0.12	0.11	0.11	0.13

Source: own calculations based on the EPH.

Table 3.2
Informality rates for salaried workers
Greater Buenos Aires, 1980, 1986, 1992 and 1998

	1980	1986	1992	1998
Total	0.20	0.24	0.32	0.38
Gender				
Men	0.18	0.20	0.29	0.36
Women	0.24	0.31	0.36	0.40
Age				
0-17	0.56	0.77	0.87	0.94
18-24	0.23	0.36	0.42	0.52
25-39	0.17	0.18	0.27	0.32
40-59	0.14	0.18	0.23	0.30
60-	0.30	0.37	0.39	0.46
Marital status				
Single	0.25	0.33	0.39	0.45
Married	0.16	0.19	0.26	0.32
Education				
Primary school drop-outs	0.30	0.36	0.44	0.63
Primary school graduates	0.23	0.30	0.39	0.47
High school drop-outs	0.19	0.27	0.41	0.41
High school graduates	0.09	0.11	0.19	0.30
College drop-outs	0.10	0.13	0.23	0.25
College graduates	0.07	0.08	0.08	0.18
Sector				
Low tech industry	0.23	0.28	0.27	0.44
Medium and high tech industry	0.12	0.16	0.27	0.29
Transport and utilities	0.16	0.13	0.23	0.39
Construction	0.33	0.49	0.58	0.65
Retail and wholesale trade	0.24	0.33	0.45	0.48
Business and financial services	0.09	0.13	0.22	0.22
Public Administration	0.01	0.02	0.04	0.20
Social services	0.07	0.10	0.10	0.15
Other services	0.47	0.46	0.63	0.66
Size of the firm (employees)				
1	0.67	0.75	0.76	0.85
From 2 to 5	0.46	0.53	0.66	0.69
From 6 to 25	0.18	0.21	0.30	0.39
From 26 to 100	0.06	0.07	0.09	0.16
From 101 to 500	0.05	0.03	0.07	0.10
More than 500	0.02	0.04	0.06	0.08
Seniority				
Until 3 months		0.64	0.66	0.76
From 3 to 6 months		0.45	0.55	0.50
From 6 months to 1 year	0.32	0.39	0.42	0.52
From 1 year to 5 years	0.20	0.24	0.35	0.38
From 5 years to 10 years	0.09	0.11	0.16	0.17
From 10 years to 20 years		0.08	0.11	0.15
More than 20 years	0.10	0.04	0.07	0.05
Weekly hours of work				
From 1 to 19	0.54	0.74	0.74	0.82
From 20 to 29	0.32	0.43	0.47	0.55
From 30 to 40	0.16	0.21	0.28	0.32
From 41 to 45	0.17	0.16	0.22	0.24
From 46 to 61	0.20	0.23	0.30	0.32
More than 62	0.20	0.23	0.35	0.39

Source: own calculations based on the EPH.

Table 4.1
Models of the informality choice for salaried workers
Greater Buenos Aires, 1980, 1986, 1992 and 1998

Year	1980			1986			1992			1998		
Model	1	2	3	1	2	3	1	2	3	1	2	3
woman	0.337	-0.120	-0.021	0.546	0.075	0.131	0.376	0.057	0.013	0.206	-0.101	-0.088
	0.073	0.092	0.103	0.060	0.075	0.083	0.064	0.080	0.094	0.052	0.066	0.078
married	-0.017	0.009	0.006	-0.079	-0.175	-0.192	0.031	0.126	0.137	-0.184	-0.079	-0.078
	0.081	0.097	0.097	0.066	0.084	0.085	0.069	0.084	0.088	0.057	0.069	0.071
age	-0.115	-0.084	-0.088	-0.148	-0.082	-0.080	-0.133	-0.092	-0.103	-0.111	-0.055	-0.069
	0.015	0.018	0.019	0.013	0.016	0.016	0.014	0.017	0.018	0.012	0.015	0.015
age squared	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
primary school graduate	-0.341	-0.244	-0.204	-0.303	-0.085	-0.055	-0.276	-0.045	0.020	-0.446	-0.307	-0.290
	0.084	0.098	0.103	0.077	0.095	0.098	0.097	0.122	0.127	0.102	0.124	0.131
high school drop-out	-0.561	-0.373	-0.246	-0.444	-0.112	-0.047	-0.371	-0.076	0.008	-0.671	-0.472	-0.404
	0.105	0.128	0.134	0.090	0.113	0.118	0.107	0.129	0.134	0.108	0.130	0.138
high school graduate	-0.885	-0.506	-0.312	-1.032	-0.567	-0.406	-0.899	-0.531	-0.374	-0.888	-0.586	-0.418
	0.135	0.155	0.170	0.113	0.132	0.141	0.116	0.137	0.143	0.112	0.136	0.146
college drop-out	-1.027	-0.916	-0.673	-0.994	-0.750	-0.574	-0.870	-0.327	-0.071	-1.175	-0.841	-0.590
	0.169	0.211	0.221	0.132	0.164	0.169	0.140	0.171	0.179	0.126	0.148	0.160
college graduate	-0.883	-0.350	-0.025	-0.987	-0.448	-0.304	-1.440	-0.760	-0.432	-1.054	-0.618	-0.287
	0.184	0.212	0.239	0.144	0.185	0.192	0.165	0.210	0.221	0.131	0.158	0.173
equivalized household income	0.000	0.069	0.075	-0.077	-0.015	0.000	-0.049	-0.036	-0.047	-0.112	0.011	0.000
	0.040	0.040	0.041	0.041	0.038	0.038	0.040	0.037	0.032	0.030	0.027	0.030
size of the firm		-0.540	-0.537		-0.456	-0.441		-0.523	-0.419		-0.467	-0.446
		0.040	0.044		0.035	0.043		0.037	0.041		0.028	0.033
seniority		-0.189	-0.181		-0.278	-0.265		-0.281	-0.279		-0.312	-0.303
		0.031	0.032		0.026	0.026		0.027	0.027		0.021	0.022
part-time job		0.632	0.765		0.733	0.808		0.441	0.586		0.715	0.853
		0.132	0.145		0.110	0.122		0.108	0.128		0.083	0.098
constant	1.605	3.218	2.760	2.505	3.361	2.953	2.514	3.944	3.573	2.760	3.750	3.605
	0.274	0.334	0.397	0.239	0.315	0.356	0.254	0.346	0.402	0.242	0.298	0.353
Observations	2457	2265	2166	3051	2670	2659	2348	1989	1989	3073	2803	2785
Sectoral dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Log likelihood	-1076	-745	-711	-1382	-882	-857	-1254	-796	-747	-1809	-1206	-1124
Pseudo R2	0.085	0.303	0.321	0.146	0.369	0.386	0.127	0.357	0.397	0.105	0.346	0.388

Note: Probit models estimated with maximum likelihood. The estimated standard deviations are shown under the coefficients. Following Maloney (1999), selection bias is not taking into account.

Figure 4.1
Probability of informality by age
Estimations based in model 3 of Table 4.1
Greater Buenos Aires, 1980, 1986, 1992 and 1998

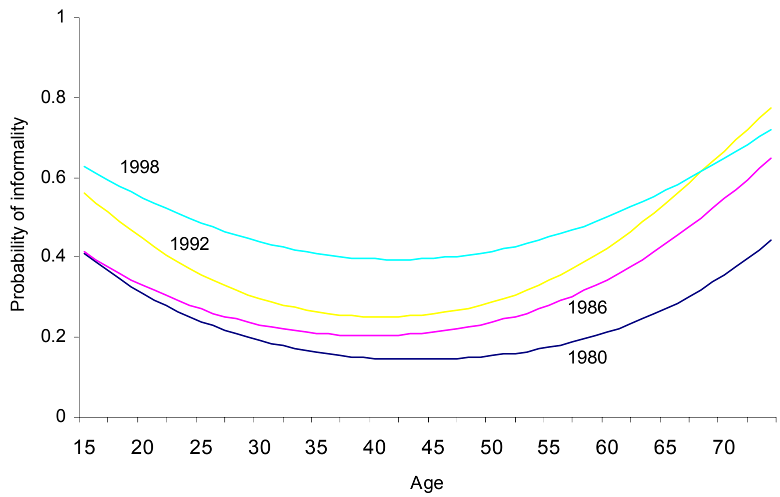
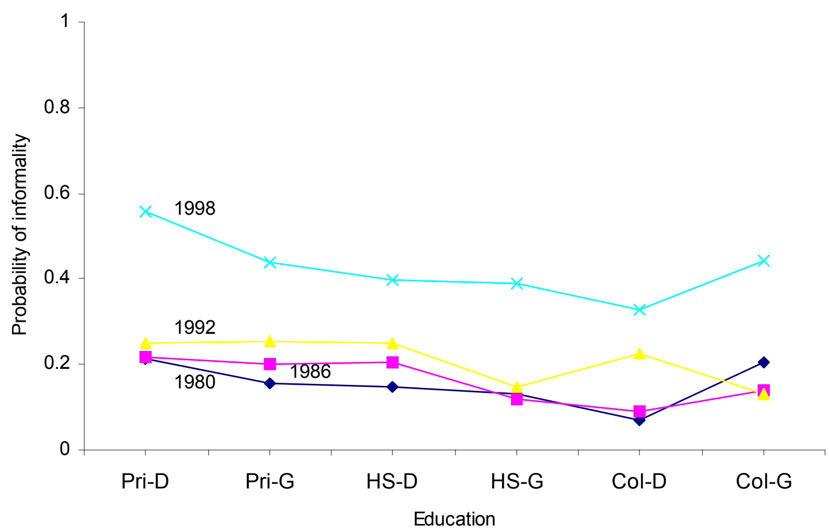
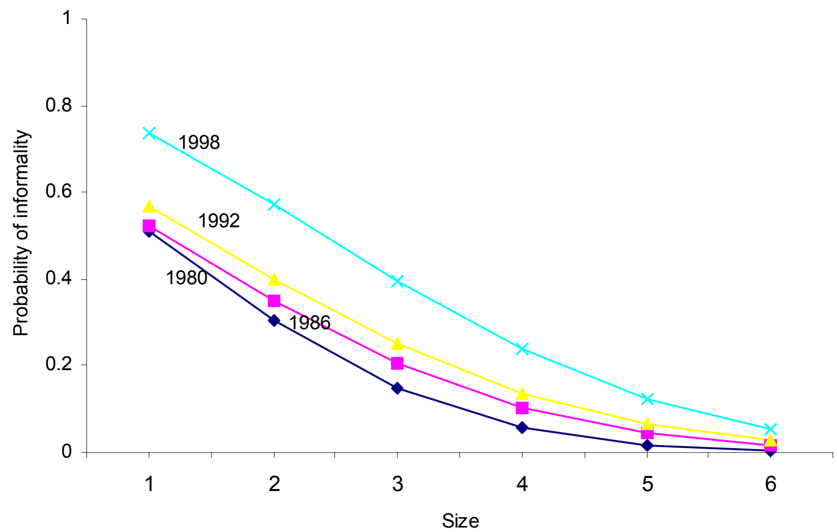


Figure 4.2
Probability of informality by education
Estimations based on model 3 of Table 4.1
Greater Buenos Aires, 1980, 1986, 1992 and 1998



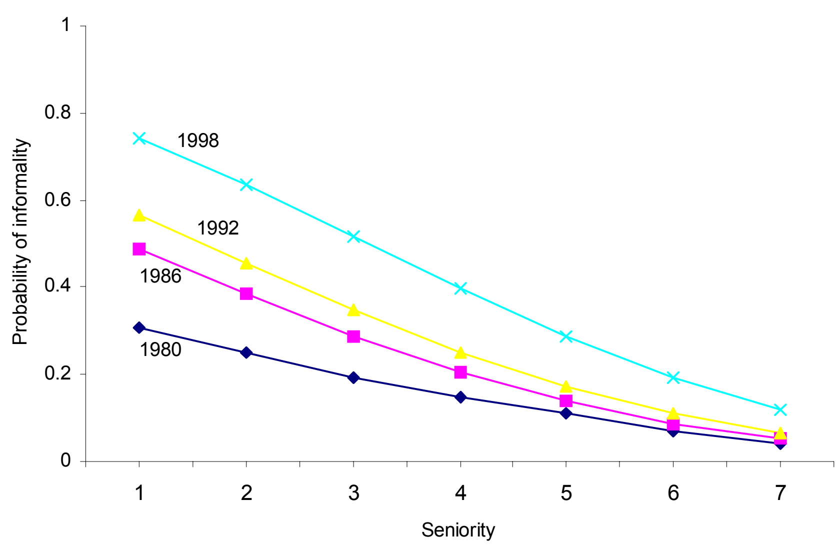
Note: Pri-D=Primary school drop-outs, Pri-G= Primary school graduates, HS-D= High school drop-outs, HS-G= High school graduates, Col-D= College drop-outs, Col-G= College graduates

Figure 4.3
Probability of informality by size of the firm
Estimations based on model 3 of Table 4.1
Greater Buenos Aires, 1980, 1986, 1992 and 1998



Note: 1=one employee, 2=from 2 to 5, 3=from 6 to 25, 4= from 26 to 100, 5=from 101 to 500 and 6=more than 500.

Figure 4.4
Probability of informality by job seniority
Estimations based on model 3 of Table 4.1
Greater Buenos Aires, 1980, 1986, 1992 and 1998



Note: 1=until 3 months, 2=from 3 to 6 months, 3=from 6 months to 1 year, 4=from 1 year to 5 years, 5=from 5 years to 10 years, 6=from 10 years to 20 years and 7=more than 20 years.

[i] The perception of other social benefits derived from the labor relationship is highly correlated with the receipt of pension: 98% of the people who declare to have a pension right, states that (s)he receives an annual bonus, takes holidays, has work insurance coverage and a right to claim severance payments if fired (GBA, 1998).

[ii] The EPH contains questions on retirement benefits only for salaried workers, and therefore they are the only group studied here. In any case, almost 75% of the workers in the Greater Buenos Aires area are wage earners, a percentage that has not substantially changed in the past two decades.

[iii] See Arango and Maloney (2000), Carpio, Klein and Novakosvski (2000) and Gasparini (2000) for more evidence on labor informality in Argentina.

[iv] This pattern is concentrated in the 1980-1992 period: in the last years large firms have slightly increased their share in employment.

[v] The model is based on Rosen (1986) and de la Rica and Lemieux (1994).

[vi] D_i is defined as the value such that $u\ (y_i\text{-}D_i)+t_i=u(y_i)$.

[vii] This probability can vary among workers. For simplicity, in the model p is kept constant.

[viii] In terms of employment, it is possible to prove that, with evasion, the number of workers is higher than in the case without evasion, but still lower than the completely non regulated case.

[ix] The rates of labor informality of Table 5.1 slightly differ from the ones reported in Figure 3.1, as they are obtained from a sample composed only by people with observations of all the explanatory variables of model 3 in Table 4.1.

[x] The residual effect is not reported, since it is negligible.

[xi] Given equation (2.9) and in order to "smooth" the results, the analysis is focused on the average of the effects (columns (v) and (vi)).

xii The South includes the cities of Comodoro Rivadavia, Río Gallegos and Tierra del Fuego (Ushuaia and Río Grande). The Northwest (NOA) includes Jujuy and Salta.