

Informality and segmentation: evidence from a self-selection model with entry barriers to formal employment in Peru

Alfredo M. Alvarado Enciso

Pontifical Catholic University of Peru (PUCP) and
Inter-American Development Bank (IDB)

The Economics of Informality Conference 2018
Bogotá, Colombia

May 28-29, 2018

Outline

- 1 Introduction
- 2 Model of segmentation and self-selection
- 3 Informality in Peru and sources
- 4 Results
- 5 Concluding remarks and research agenda

The problem of informality

- In Peru, 3 out of 4 workers are informal (INEI, 2014)

The problem of informality

- In Peru, 3 out of 4 workers are informal (INEI, 2014)
- Informality reduces growth (Loayza, 2008)

The problem of informality

- In Peru, 3 out of 4 workers are informal (INEI, 2014)
- Informality reduces growth (Loayza, 2008)
- Informal employment is related to low income, poverty and exclusion (Perry et. al, 2007)

The problem of informality

- In Peru, 3 out of 4 workers are informal (INEI, 2014)
- Informality reduces growth (Loayza, 2008)
- Informal employment is related to low income, poverty and exclusion (Perry et. al, 2007)
- Labor informality has been a topic of debate in public policies in Peru: min. wages, special tax regimes, special labor regimes, etc

The problem of informality

- In Peru, 3 out of 4 workers are informal (INEI, 2014)
- Informality reduces growth (Loayza, 2008)
- Informal employment is related to low income, poverty and exclusion (Perry et. al, 2007)
- Labor informality has been a topic of debate in public policies in Peru: min. wages, special tax regimes, special labor regimes, etc
- Is informality a problem?

Debate about the nature of labor informality

- The academic debate has revolved around whether informality indicates:

Debate about the nature of labor informality

- The academic debate has revolved around whether informality indicates:
 - ▶ Segmentation in the labor market (Fei & Ranis, 1961; Lewis, 1963; Harris & Todaro, 1970; Doeringer & Piore, 1971)

Debate about the nature of labor informality

- The academic debate has revolved around whether informality indicates:
 - ▶ Segmentation in the labor market (Fei & Ranis, 1961; Lewis, 1963; Harris & Todaro, 1970; Doeringer & Piore, 1971)
 - ▶ A competitive and voluntary option (De Soto, 1986; Maloney, 2004; Perry, y otros, 2007)

Debate about the nature of labor informality

- The academic debate has revolved around whether informality indicates:
 - ▶ Segmentation in the labor market (Fei & Ranis, 1961; Lewis, 1963; Harris & Todaro, 1970; Doeringer & Piore, 1971)
 - ▶ A competitive and voluntary option (De Soto, 1986; Maloney, 2004; Perry, y otros, 2007)
 - ▶ A mixed of both (Fields, 2005; Chen, y otros, 2005; Bacchetta & Ernst, 2009)

Methodological background

- Dickens y Lang (1985) use a switching regression model with unknown regimes: a wage equation with human capital variables and another without them.

Methodological background

- Dickens y Lang (1985) use a switching regression model with unknown regimes: a wage equation with human capital variables and another without them.
- Magnac (1991) uses a multivariate generalized Tobit with three dependent variables: labor force participation, choice of sectors (self-employed or employee/blue collar with entry cost) and wage.

Methodological background

- Dickens y Lang (1985) use a switching regression model with unknown regimes: a wage equation with human capital variables and another without them.
- Magnac (1991) uses a multivariate generalized Tobit with three dependent variables: labor force participation, choice of sectors (self-employed or employee/blue collar with entry cost) and wage.
- Yamada (1994) uses structural probits and mincer earning equations corrected by selection bias of the employment sector (self-employed, informal and formal salaried).

Methodological background

- Dickens y Lang (1985) use a switching regression model with unknown regimes: a wage equation with human capital variables and another without them.
- Magnac (1991) uses a multivariate generalized Tobit with three dependent variables: labor force participation, choice of sectors (self-employed or employee/blue collar with entry cost) and wage.
- Yamada (1994) uses structural probits and mincer earning equations corrected by selection bias of the employment sector (self-employed, informal and formal salaried).
- Guindling (1991) tests whether there are differences of human capital returns among self-employed/salaried formal/informal workers.

Heckman y Hotz (1986)

- The incentive structure of the workers focuses on maximizing the utility function rather than the income function.
⇒ Workers also choose their sector based on a comparison of nonpecuniary concerns.

Heckman y Hotz (1986)

- The incentive structure of the workers focuses on maximizing the utility function rather than the income function.
 - ⇒ Workers also choose their sector based on a comparison of nonpecuniary concerns.
- Previous studies assume that competitive and non-competitive segments are exogenous.
 - ⇒ Self-employed sector is not always competitive, and salaried sector is not always segmented.

Methodological background II

- Cunningham y Maloney (2001), through factor and cluster analysis, determines the proportion of small informal enterprises was competitive and segmented.

Methodological background II

- Cunningham y Maloney (2001), through factor and cluster analysis, determines the proportion of small informal enterprises was competitive and segmented.
- Pratap y Quitin (2006), through non-parametric techniques (PSM and D&D), assesses the segmentation hypothesis by comparing informal workers who earns the same as formal workers.

Methodological background II

- Cunningham y Maloney (2001), through factor and cluster analysis, determines the proportion of small informal enterprises was competitive and segmented.
- Pratap y Quitin (2006), through non-parametric techniques (PSM and D&D), assesses the segmentation hypothesis by comparing informal workers who earns the same as formal workers.
- Günther y Launov (2012), through finite mixture model with sample selection, determine the number of unobservable segments of informal sector endogenously and calculates the % of involuntary informal workers.

Methodological background II

- Cunningham y Maloney (2001), through factor and cluster analysis, determines the proportion of small informal enterprises was competitive and segmented.
- Pratap y Quitin (2006), through non-parametric techniques (PSM and D&D), assesses the segmentation hypothesis by comparing informal workers who earns the same as formal workers.
- Günther y Launov (2012), through finite mixture model with sample selection, determine the number of unobservable segments of informal sector endogenously and calculates the % of involuntary informal workers.
- Alcaraz, Chiquiar y Salcedo (2015), through a utility maximization model with entry barriers and self-selection, determine the % of involuntary informal workers.

Review II

Günther y Launov (2012)

Pratap y Quitin (2006) do not consider the heterogenous nature of informal labor market and Cunningham y Maloney (2001) does not take into account the sample selection.

Alcaraz, Chiquiar y Salcedo (2015)

The finite mixture model (Günther y Launov, 2012) does correct the sample selection of labor participation but does not address the issue of formal/informal selection.

Evidence in Peru

- Yamada (1994) self-employed workers have competitive earnings and represent a voluntary option while informal salaried are segmented in Lima in 1985, 1986 and 1990.
- Palomino (2011), replicated Prata y Quitin (2006)'s model. According to him, around 50 % of informal workers are segmented in Lima in 2003.
- Baldárrago (2015) replicated Guindling (1991)'s methodology. According to her, self-employed represent a competitive labor market while salaried are segmented in the south of Peru in 2013.
- Tello (2015) replicated Günther y Launov (2012)'s methodology. According to him, 73 % of informal are involuntary in Peru in 2014.

Objetives

- To adapt Alcaraz, Chiquiar y Salcedo (2015)'s methodology to the Peruvian case
- To extend the model allowing correction for sample selection of labor participation
- To test whether there are segmented, competitive or a mixed of both labor market
- To calculate the proportion of involuntary informal workers
- To propose a research agenda

Outline

- 1 Introduction
- 2 Model of segmentation and self-selection**
- 3 Informality in Peru and sources
- 4 Results
- 5 Concluding remarks and research agenda

Assumptions of the model

- The individual maximizes his utility and not the income function (non-monetary valuations)

Assumptions of the model

- The individual maximizes his utility and not the income function (non-monetary valuations)
- The decision to join the labor force is considered (we include those who do not work)

Assumptions of the model

- The individual maximizes his utility and not the income function (non-monetary valuations)
- The decision to join the labor force is considered (we include those who do not work)
- Self-selection in the informal/formal labor market is considered

Assumptions of the model

- The individual maximizes his utility and not the income function (non-monetary valuations)
- The decision to join the labor force is considered (we include those who do not work)
- Self-selection in the informal/formal labor market is considered
- There are entry barriers to formal employment

Model

Labor force participation equation:

$$L^* = Z_i\alpha + \varepsilon_{i,1} \quad (1)$$

Choice of the segment formal/informal:

$$F^* = X_i\beta + \varepsilon_{i,2} \quad (2)$$

$$\begin{bmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \end{bmatrix} | Z, X \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

Choice of the segments

$$\left\{ \begin{array}{l} \text{do not work if } L^* < 0 \\ \text{formal if } L^* > 0 \text{ and } F^* > 0 \text{ and is hired} \\ \text{informal if } L^* > 0 \text{ and } F^* < 0 \text{ or } F^* > 0 \text{ and is not hired} \end{array} \right.$$

Model II

We define the hiring parameter δ as the probability of being hired or access to formal employment. Therefore, the probability of choosing an employment is:

Probability of not working

$$p(\text{do not work}) = P(L^* > 0) = \Theta(Z_{i,1}\alpha) \quad (3)$$

Model II

We define the hiring parameter δ as the probability of being hired or access to formal employment. Therefore, the probability of choosing an employment is:

Probability of not working

$$p(\text{do not work}) = P(L^* > 0) = \Theta(Z_{i,1}\alpha) \quad (3)$$

Probability of working in a formal employment

$$\begin{aligned} p(\text{formal}) &= P(F^* > 0 | L^* > 0) P(L^* > 0) \delta \\ &= \delta P(X_{i,2}\beta > 0 | Z_{i,1}\alpha > 0) P(Z_{i,1}\alpha > 0) \end{aligned} \quad (4)$$

Model II

We define the hiring parameter δ as the probability of being hired or access to formal employment. Therefore, the probability of choosing an employment is:

Probability of not working

$$p(\text{do not work}) = P(L^* > 0) = \Theta(Z_{i,1}\alpha) \quad (3)$$

Probability of working in a formal employment

$$\begin{aligned} p(\text{formal}) &= P(F^* > 0 | L^* > 0) P(L^* > 0) \delta \\ &= \delta P(X_{i,2}\beta > 0 | Z_{i,1}\alpha > 0) P(Z_{i,1}\alpha > 0) \\ &= \delta \text{BivariateNormal}(X_{i,2}\beta, Z_{i,1}\alpha, \rho) \end{aligned} \quad (4)$$

Model III

Probability of working in an informal employment

$$\begin{aligned} p(\text{informal}) &= P(F^* < 0 | L^* > 0)P(L^* > 0) \\ &\quad + (1 - \delta)P(F^* > 0 | L^* > 0)P(L^* > 0) \end{aligned}$$

(5)

Model III

Probability of working in an informal employment

$$\begin{aligned} p(\text{informal}) &= P(F^* < 0 | L^* > 0)P(L^* > 0) \\ &+ (1 - \delta)P(F^* > 0 | L^* > 0)P(L^* > 0) \\ &= P(X_{i,2}\beta < 0 | Z_{i,1}\beta > 0)P(Z_{i,1}\beta > 0) \\ &+ (1 - \delta)P(X_{i,2}\beta > 0 | Z_{i,1}\beta > 0)P(Z_{i,1}\beta > 0) \end{aligned} \tag{5}$$

Probability of working in an informal employment

$$\begin{aligned} p(\text{informal}) &= P(F^* < 0 | L^* > 0)P(L^* > 0) \\ &+ (1 - \delta)P(F^* > 0 | L^* > 0)P(L^* > 0) \\ &= P(X_{i,2}\beta < 0 | Z_{i,1}\beta > 0)P(Z_{i,1}\beta > 0) \\ &+ (1 - \delta)P(X_{i,2}\beta > 0 | Z_{i,1}\beta > 0)P(Z_{i,1}\beta > 0) \\ &= \text{BivariateNormal}(-X_{i,2}\beta, Z_{i,1}\alpha, -\rho) \\ &+ (1 - \delta)\text{BivariateNormal}(X_{i,2}\beta, Z_{i,1}\alpha, \rho) \end{aligned} \tag{5}$$

Model IV

So log-likelihood function and the parameter constrain is:

$$\begin{aligned} & \sum_{\text{formal}} [\ln(\delta BVN(X_{i,2}\beta, Z_{i,1}\alpha, \rho))] \\ + & \sum_{\text{informal}} [\ln(BivariateNormal(-X_{i,2}\beta, Z_{i,1}\alpha, -\rho) \\ & + (1 - \delta)BivariateNormal(X_{i,2}\beta, Z_{i,1}\alpha, \rho))] \\ + & \sum_{\text{do not work}} [\ln(\Theta(Z_{i,1}\alpha))], \text{ where } 0 < \delta < 1 \end{aligned} \tag{6}$$

Modelo V

We use the hiring parameter δ to calculate the proportion of involuntary informal workers. Let FO be the number of formal workers, I the number of informal workers and M the number of workers that would prefer to be formal ($FO = \delta M$). Therefore,

Proportion of involuntary informal workers

$$\frac{\text{involuntary informal workers}}{\text{total informal workers}} \% = \frac{(1 - \delta)M}{I} = \frac{(1 - \delta)FO}{\delta I} \quad (7)$$

Outline

- 1 Introduction
- 2 Model of segmentation and self-selection
- 3 Informality in Peru and sources**
- 4 Results
- 5 Concluding remarks and research agenda

Definition of informality

- ILO (1993, 2003), Delhi Group 1997 and the last version of the SNA (2008).
- A Satellite Account of the Informal Economy in Peru (INEI, 2014)
- Identification of salaried informal workers with ENAHO (official from 2012):
 - ▶ Employees without health insurance granted by their employers or in a unregistered firm
- Identification of salaried informal workers with ENAHO (ad hoc narrow definition):
 - ▶ Employees without health insurance granted by their employers, without payment to pension insurance, without contract, in an unregistered firm, the firm do not have books of account

Stylized facts

- During 2012-2016, the informality rate (of urban salaried from private sector) has a slight negative tendency.
- Informality stopped shrinking in the narrow def. because of the reduction in the growth in 2014 (Cespedes, 2015).

Evolution in the informality rate and growth of GDP, 2012-2016

Years	Informal (official)	Informal (narrow def.)	Economic growth
2012	63%	84%	6.1%
2013	63%	85%	5.9%
2014	62%	84%	2.4%
2015	60%	85%	3.3%
2016	59%	85%	3.9%

Descriptive statistics of the model variables

Variables	Informal (official)	Informal (narrow def.)
<i>Education</i>		
Primary Or Less	85%	96%
Secondary	70%	89%
Non-University Higher Education	48%	72%
University Higher Education	42%	72%
Years old (mean)	32.51	35.44
Married	41%	39%
Head of HH	29%	22%
Has any insurance different from job's	43%	52%
White collars	38%	16%
Work more than 35 hours	63%	84%
<i>Geographic area</i>		
Costa Norte	70%	89%
Costa Centro	62%	84%
Costa Sur	71%	88%
Sierra Norte	76%	92%
Sierra Centro	75%	90%
Sierra Sur	70%	88%
Selva	82%	93%
Lima Metropolitana	51%	80%

Outline

- 1 Introduction
- 2 Model of segmentation and self-selection
- 3 Informality in Peru and sources
- 4 Results**
- 5 Concluding remarks and research agenda

Results

- The % of involuntary informal workers is 10 % or 5 % depending of the definition.
- The low levels are coherent. In the Mexican case, the % is 15 (Alcaraz, Chiquiar, & Salcedo, 2012).
- Estimates from Tello (2015) are between 11 % and 73 % depending of the definition.

Variables	Informal (official)		Informal (narrow)	
	(1)	(2)	(3)	(4)
Years old	-0.133*** (-41.73)	-0.112*** (-31.35)	-0.135*** (-37.65)	-0.119*** (-29.01)
Years old ²	0.00140*** (38.83)	0.00115*** (27.65)	0.00144*** (35.41)	0.00126*** (26.33)
Primary or less	-1.98e-13 (.)	-3.67e-14 (.)	9.46e-15 (.)	1.49e-14 (.)
Secondary	-0.610*** (-28.06)	-0.707*** (-26.69)	-0.592*** (-25.09)	-0.692*** (-23.37)
Non-university higher education	-1.216*** (-41.16)	-1.386*** (-41.15)	-1.197*** (-36.90)	-1.392*** (-36.20)
University higher education	-1.392*** (-43.80)	-1.623*** (-45.13)	-1.388*** (-39.39)	-1.663*** (-39.96)
No married	(.)	-2.39e-16 (.)	(.)	-1.67e-16 (.)
Married	(.)	-0.414*** (-23.73)	(.)	-0.413*** (-21.15)
No head of HH	(.)	-3.05e-16 (.)	(.)	2.43e-16 (.)
Head of HH	(.)	-0.134*** (-7.48)	(.)	-0.146*** (-7.26)
Has health insurance diff. of job's	(.)	2.005*** (67.22)	(.)	1.884*** (55.44)
Constant	3.367*** (45.49)	3.564*** (45.22)	3.432*** (42.57)	3.673*** (40.57)
p	0.516*** (26.37)	0.0559 (1.82)	0.535*** (25.63)	0.126*** (3.76)
$\hat{\sigma}$	1.826*** (14.25)	2.092*** (21.79)	1.250*** (11.44)	1.261*** (18.65)
%involuntary informal	8.94*** (-0.012)	6.85*** (-0.007)	4.41*** (-0.005)	4.36*** (-0.003)
Geo. areas and years fixed eff.	Yes	Yes	Yes	Yes
Observations	158520	158520	158520	158520

Robustness check

- There are more involuntary informal white collars. Unlike blue collars, white collars really care about being formal.
- There are less involuntary informal full-time workers. Part-time workers really care about being formal.

	Informal (official)		Informal (narrow)	
	(1)	(2)	(3)	(4)
Benchmark	8.94*** (-0.012)	6.85*** (-0.007)	4.41*** (-0.005)	4.36*** (-0.003)
White collars	10.46*** (-0.013)	8.67*** (-0.008)	4.42*** (-0.004)	4.27*** (-0.003)
Work more than 35 hours	5.8*** (-0.016)	7.04*** (-0.008)	1.99*** (-0.004)	2.69*** (-0.002)

Outline

- 1 Introduction
- 2 Model of segmentation and self-selection
- 3 Informality in Peru and sources
- 4 Results
- 5 Concluding remarks and research agenda

Concluding remarks

- The hypothesis of the existence of multi-segmented labor markets is confirmed
- Around 10 % of informal workers are involuntary
- Limitations of the study:
 - ▶ The model has strong assumptions about the distributions
 - ▶ There might be endogeneity with the variables correlated with earning

Research agenda

- Correct the possible endogeneity of the variables correlated with the income (IV, control function, etc).
- Explore the public policy effects over the hiring parameter within a model of self-selection of informal employment.

Thanks