

# Housing informality and labor informality in space

## In search of the missing links

**Gustavo A. García**

School of Finance, Economics and Government

Universidad EAFIT

[ggarci24@eafit.edu.co](mailto:ggarci24@eafit.edu.co)

**Erika R. Badillo**

Faculty of Economics

Universidad Autónoma Latinoamericana

[erika.badilloen@unaula.edu.co](mailto:erika.badilloen@unaula.edu.co)

**Juan Manuel Aristizábal**

Faculty of Accounting, Economic and Administrative Sciences

Universidad de Manizales

[jm.aristizabal@umanizales.edu.co](mailto:jm.aristizabal@umanizales.edu.co)

20th IEA World Congress

Medellín - Colombia

11-15 December 2023

# Outline

- Motivation
- Contribution
- Relevant literature
- Data and descriptive evidence
- Econometric model
- Results
- Conclusions

# Motivation

- Many developing countries are characterized by the prevalence of **housing informality, labor informality** and **informality of transport**
- This triple informality in housing, employment and transport is an important factor driving the low levels in productivity and wellbeing observed in many cities in developing countries (Vargas et al., 2017; OECD, 2018; Ferreyra and Roberts, 2018; Azunre et al, 2021)
- Housing and labor informality are the phenomena with **the highest incidence** at the urban level and they are closely related to socioeconomic indicators such as poverty, inequality, quality of life and, in general, **the development of cities and their productivity** (UN-Habitat, 2016):

	Housing informality	Labor informality
World	32 %	61 %
Latin America	25 %	50 %
Colombia	22 %	50 %

Source: UN-Habitat (2016), ILO (2018)

# Motivation

- The cities in developing countries are characterized by the existent of areas hosting low-income workers living in precarious conditions
- These lower incomes might induce people to live in informal settlements, so that economic vulnerability and precarious working conditions may be within the causes of the emergence of informal housing (Lombard and Mitlin, 2020)
- The precarious working conditions along with higher restrictions to access to the financial system that face informal workers are thought to be important mechanisms that explain the spatial choice of housing and its characteristics (Vargas et al., 2017)

# Contribution

- Despite the extensive literature on both housing informality and labor informality, there is a lack of evidence on the relationship between these two phenomena in space
- This paper aims to offer new evidence on the links between precarious housing conditions (informal housing) and limited access to quality employment (labor informality), considering the relationship and spatial dimension of these two phenomena
- Using data at the intra-urban level in Medellin (Colombia), we identified the effect of labor informality on housing informality through spatial simultaneous equation models, which take into account the relationship between informal housing and labor informality and the spatial dependence

# Relevant literature

Several studies focus on analyzing the determinants of housing informality and labor informality separately, forgetting the interdependence between them

Among the few studies that simultaneously analyze these phenomena, they agree that precarious working conditions may be within the causes of precarious housing conditions

## Determinants of housing informality

- UN-Habitat (2013): lag in infrastructure
- Inostroza et al. (2013): urban growth
- Bonet et al. (2016): labor informality
- Posada and Moreno (2019): labor informality, structure of the cities and price of land
- Posada (2018): population growth

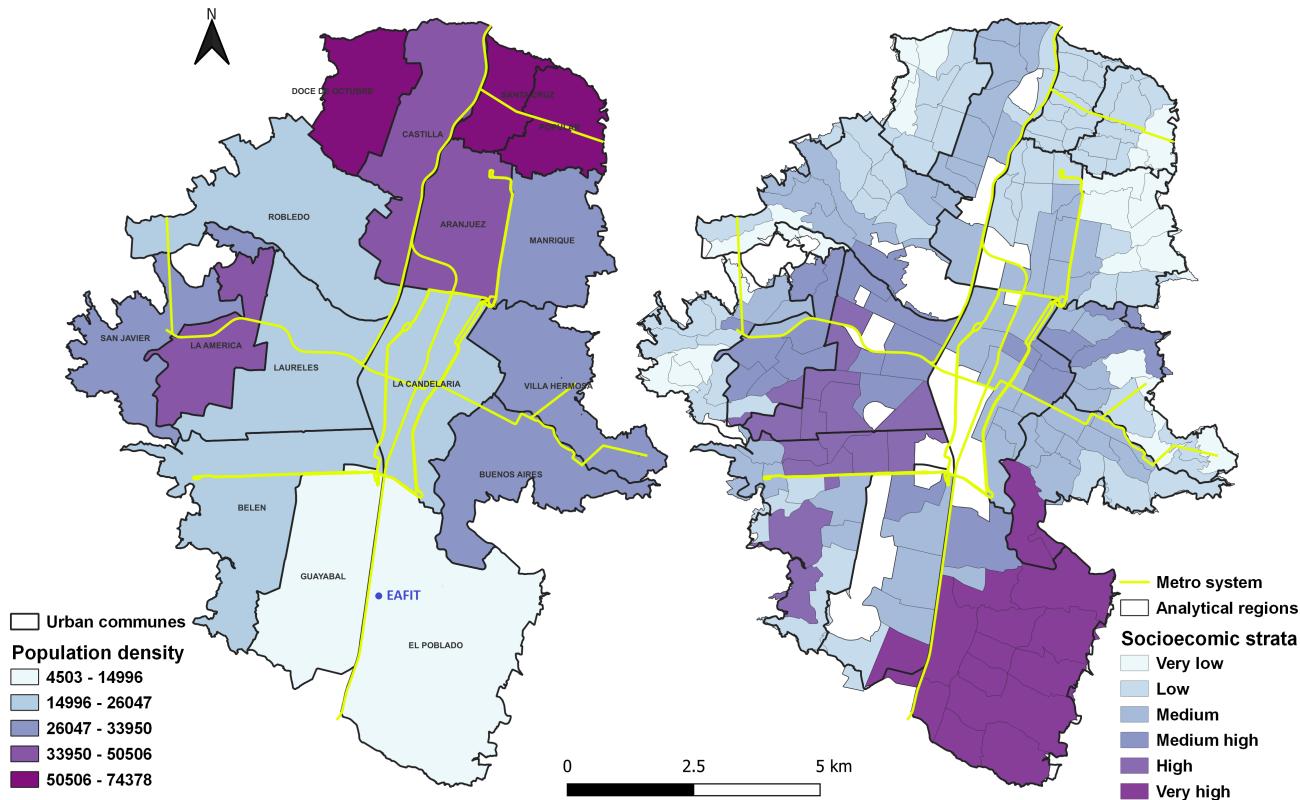
## Determinants of labor informality

- Boisjoly et al. (2017): public transport
- Del Carpio and Patrick (2020): growth of cities
- Gutierrez (2021): inequality
- Lombard et al. (2021): level of income

# Data and descriptive evidence

- Population: 2.6 millions (Colombia: 48 millions, Bogotá: 8 millions)
- Area: 380  $km^2$  (Colombia: 1,141,748  $km^2$ )
- Density: 6749  $km^2$  (Bogotá: 4531  $km^2$ )
- Unemployment: 8.6% (Colombia: 9.2%)
- Housing informality: 7% (Colombia: 22%)
- Labor informality: 44% (Colombia: 50%)
- Administrative division: 16 communes and 243 neighborhoods

Figure 1. Study area: Medellín



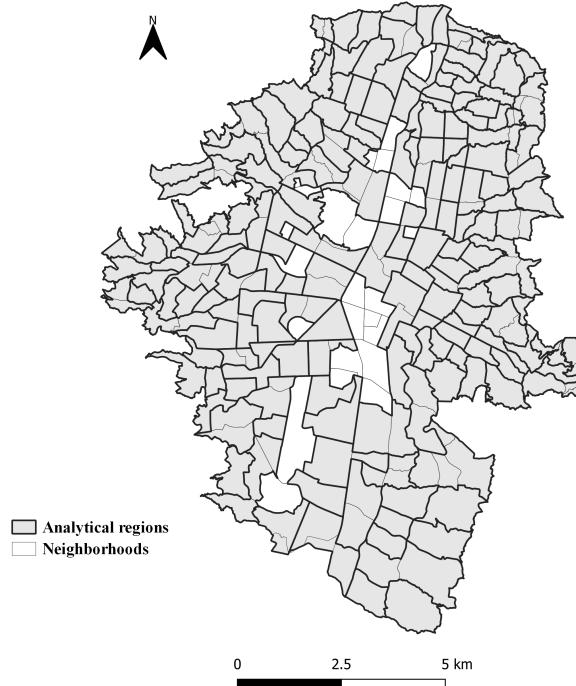
# Data and descriptive evidence

- The data used in this paper came from [the Quality of Life Survey \(QLS\)](#) for Medellín in 2017
- This cross-section survey contains individual-level information on household characteristics, demographics, education, social security, labor market, and indicators of poverty and socioeconomic conditions
- The sample includes a total of [42806](#) individuals ([12205](#) households)

# Data and descriptive evidence

- Our spatial unit of analysis are the analytical regions, which are homogeneous regions that are defined in terms of (Duque et al., 2012):
  - neighborhood limits
  - socioeconomic characteristics
  - each unit contains at least 100 household
- These analytical regions are calculated using the Max-p-regions technique (Duque et al., 2012), which is a technique to aggregation of areas into a maximum number of homogeneous regions that guarantees statistical representativeness and minimizes intraregional heterogeneity
- We have **176 urban analytical regions** (average area:  $0.521 \text{ km}^2$ )

**Figure 2. Analytic regions (176) and neighborhoods (243)**



# Data and descriptive evidence

- **Housing informality**

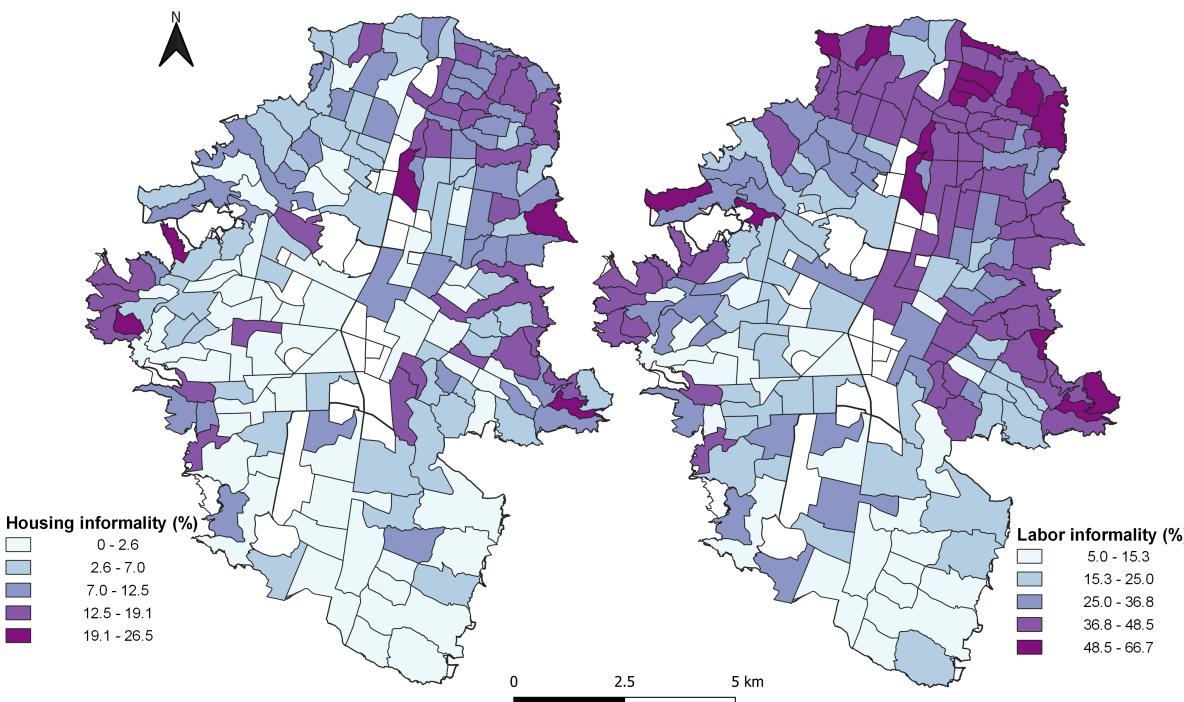
A dwelling is considered informal if it has at least one of the following conditions (UN-Habitat, 2003):

- walls that are not made of durable materials
- overcrowding
- lack of aqueduct
- toilet not connected to the aqueduct

- **Labor informality**

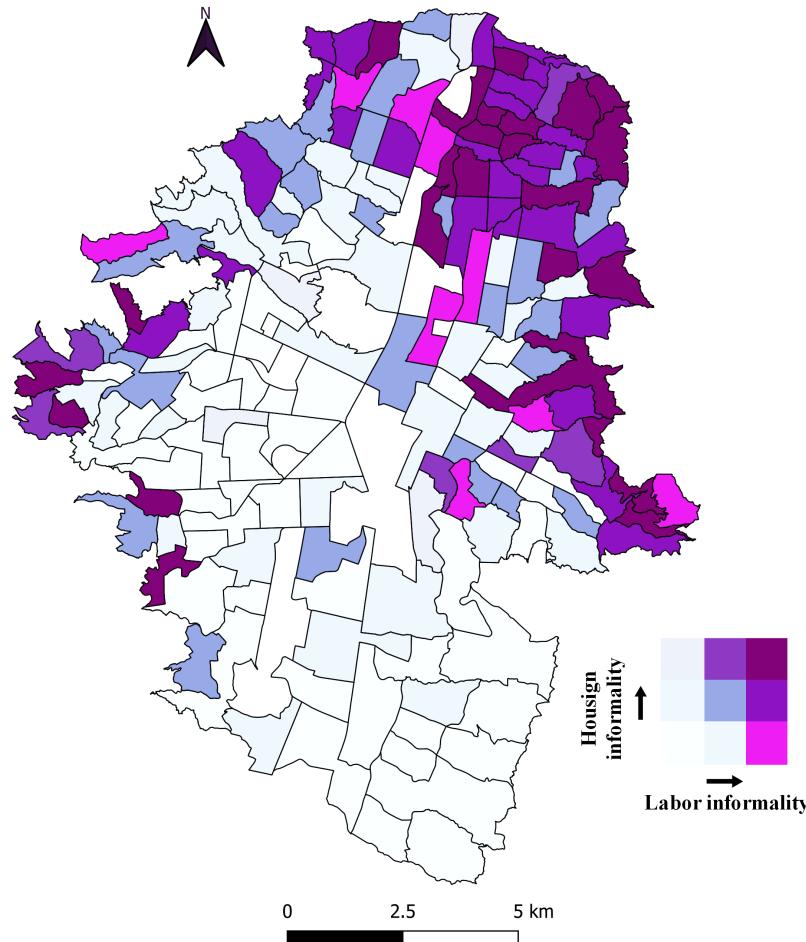
Informal workers are those employees who do not contribute to the social security system: health and pension

**Figure 3. Housing informality and labor informality**



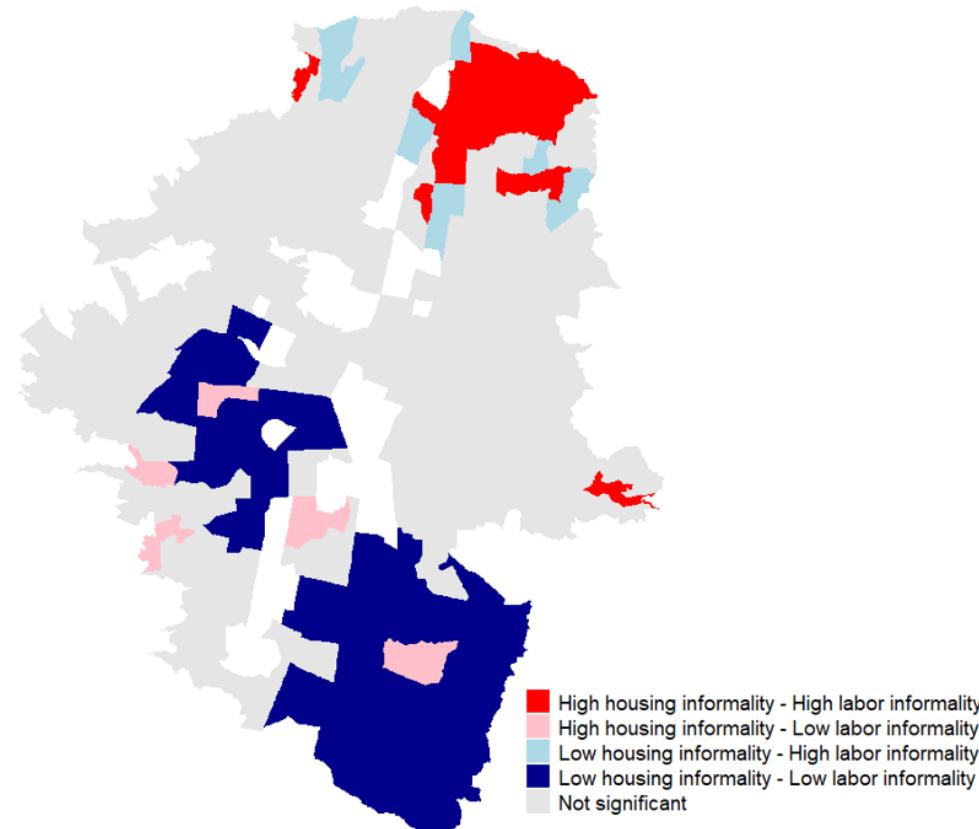
# Data and descriptive evidence

Figure 4. Bivariate map of housing informality and labor informality



# Data and descriptive evidence

Figure 5. Bivariate LISA map of housing informality and labor informality



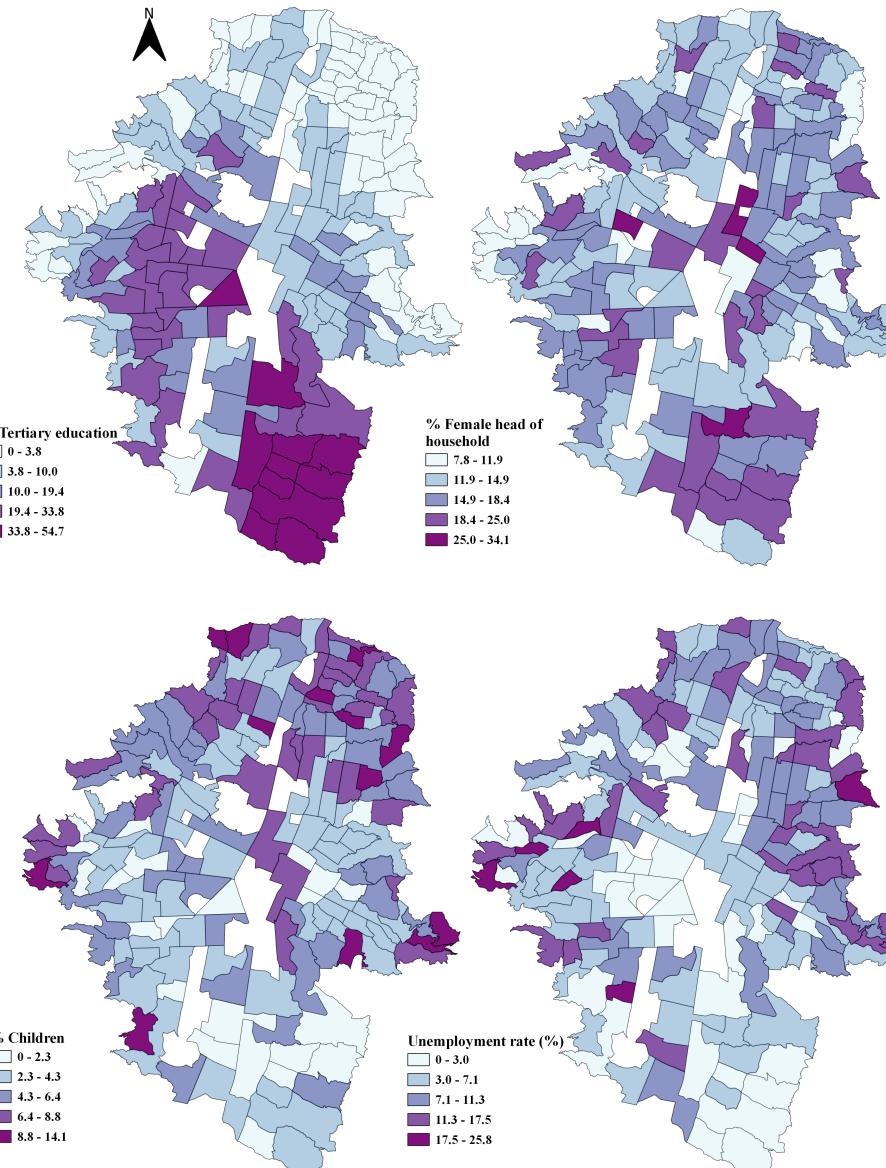
Note: number of permutations = 999 and significance level at 5%. First-order

Queen criterion of contiguity was used to calculate the spatial weights

- Regions with high informal housing present simultaneously high levels of labor informality, in particular in the east, west and north of the city  $\implies$  urban informality is located in poor peripheral areas and there is a spatial simultaneity between types of informality
- Housing informality and labor informality are phenomena with a **high spatial persistence**

# Data and descriptive evidence

Figure 6. Spatial distribution of the control variables



# Data and descriptive evidence

Table 1. Descriptive statistics of informality, control and instrumental variables

	Mean	SD	P25	Median	P75	Min	Max
Housing informality (%)	6.83	6.11	1.93	5.74	10.79	0.00	26.45
Labor informality (%)	33.40	14.01	20.06	35.56	44.80	4.87	64.00
Control variables							
% Tertiary education	11.39	12.02	2.41	6.37	16.87	0.00	54.73
% Female head of household	15.63	3.95	12.90	15.30	17.58	7.79	34.14
% Children	5.02	2.58	3.12	4.69	6.83	0.00	14.13
Unemployment rate (%)	7.84	5.17	3.81	6.80	10.95	0.00	25.79
Population density	24.59	19.30	11.05	18.79	32.76	2.11	107.84
Instrumental variables							
Housing informality in 2011 (%)	8.40	9.38	0.00	5.90	12.26	0.00	51.72
Euclidean distance to CBD (km)	2.15	0.92	1.55	2.15	2.86	0.00	4.79

Note: Population density = population/km<sup>2</sup>

## Econometric model

To measure the relationship between housing informality and labor informality while considering the simultaneity of these two phenomena, we estimate the following system of two equations:

$$\text{housing informality}_r = \beta_{11} \text{labor informality}_r + \mathbf{x}_r \boldsymbol{\omega}_1 + u_{1r} \quad (1)$$

$$\text{labor informality}_r = \beta_{21} \text{housing informality}_r + \mathbf{x}_r \boldsymbol{\omega}_2 + u_{2r}, \quad (2)$$

Term  $\mathbf{x}_r$  is a vector of control variables, which includes population with tertiary education, the share of female heads of household, the share of children under six years of age, population density, and the unemployment rate

Our coefficients of interest to estimate are  $\beta_{11}$  and  $\beta_{21}$ , which represent the effect of labor informality on informal housing and vice versa, respectively

# Econometric model

To consider the spatial dimension of housing informality and labor informality, we estimate four types of spatial simultaneous models:

## Spatial Lag Model (SLM)

$$\text{housing informality}_r = \beta_{11} \text{labor informality}_r + \rho_1 W \text{housing informality}_r + \mathbf{x}_r \boldsymbol{\omega}_1 + u_{1r} \quad (3)$$

$$\text{labor informality}_r = \beta_{21} \text{housing informality}_r + \rho_2 W \text{labor informality}_r + \mathbf{x}_r \boldsymbol{\omega}_2 + u_{2r} \quad (4)$$

## Spatial Error Model (SEM)

$$\text{housing informality}_r = \beta_{11} \text{labor informality}_r + \mathbf{x}_r \boldsymbol{\omega}_1 + u_{1r} \quad (5)$$

$$u_{1r} = \gamma_1 W u_{1r} + \varepsilon_{1r} \quad (6)$$

$$\text{labor informality}_r = \beta_{21} \text{housing informality}_r + \mathbf{x}_r \boldsymbol{\omega}_2 + u_{2r} \quad (7)$$

$$u_{2r} = \gamma_2 W u_{2r} + \varepsilon_{2r}, \quad (8)$$

Where  $W$  is a matrix ( $r \times r$ ) of spatial connectivity among analytical regions, which, in our case, is a standardized first-order Queen type

# Econometric model

## Spatial Autoregressive Combined Model (SAC)

$$\begin{aligned} \text{housing informality}_r &= \beta_{11} \text{labor informality}_r + \rho_1 W \text{housing informality}_r \\ &\quad + \mathbf{x}_r \boldsymbol{\omega}_1 + u_{1r} \end{aligned} \tag{9}$$

$$u_{1r} = \gamma_1 W u_{1r} + \varepsilon_{1r} \tag{10}$$

$$\begin{aligned} \text{labor informality}_r &= \beta_{21} \text{housing informality}_r + \rho_2 W \text{labor informality}_r \\ &\quad + \mathbf{x}_r \boldsymbol{\omega}_2 + u_{2r} \end{aligned} \tag{11}$$

$$u_{2r} = \gamma_2 W u_{2r} + \varepsilon_{2r} \tag{12}$$

## Spatial Durbin Model (SDM)

$$\begin{aligned} \text{housing informality}_r &= \beta_{11} \text{labor informality}_r + \lambda_1 W \text{labor informality}_r \\ &\quad + \rho_1 W \text{housing informality}_r \\ &\quad + \mathbf{x}_r \boldsymbol{\omega}_1 + W \mathbf{x}_r \boldsymbol{\theta}_1 \\ &\quad + u_{1r} \end{aligned} \tag{13}$$

$$\begin{aligned} \text{labor informality}_r &= \beta_{21} \text{housing informality}_r + \lambda_2 W \text{housing informality}_r \\ &\quad + \rho_2 W \text{labor informality}_r \\ &\quad + \mathbf{x}_r \boldsymbol{\omega}_2 + W \mathbf{x}_r \boldsymbol{\theta}_1 \\ &\quad + u_{2r} \end{aligned} \tag{14}$$

# Econometric model

Estimating the proposed system of equations may face two problems related to [endogeneity](#) (Le Gallo and Fingleton, 2021):

- including endogenous variables as regressors, that is, the inclusion of the dependent variable from the second equation (labor informality) in the first equation and the inclusion of the dependent variable from the first equation (housing informality) in the second equation
- including a spatial lag in the model

To deal with these endogeneity problems, we use an [Instrumental Variable \(IV\) approach](#):

- we use the Euclidean distance (in km) from the centroids of each analytical region to Medellín's Central Business District (CBD) as an instrument for the labor informality variable  $\implies$  [spatial mismatch hypothesis](#)
- we use the level of housing informality lagged in time (2011) as an instrument for the current housing informality  $\implies$  [temporal inertia of housing informality](#)
- endogeneity problem associated with the inclusion of a spatial lag in the models is corrected including spatial lags of a superior order of control variables (Kelejian and Robinson, 1993; Kelejian and Prucha, 1998; Fingleton and Le Gallo, 2008).  $\implies$  [spatial lags of order 2](#) (López et al., 2020)

Another methodological aspect to be considered in estimating the proposed system of spatial simultaneous equations is [the presence of cross-equation correlations in the error terms](#)  $\implies$  [informal transport](#)

# Econometric model

- In summary, our model is a system of spatial simultaneous equations that allows for correlation between the errors of the equations
- We use the Generalized Spatial Three-Stage Least Square (GS3SLS) estimator proposed by Kelejian and Prucha (2004) to estimate this model
- This procedure first corrects the endogeneity problem using a two-stage estimator, namely the GS2SLS method. Then, it considers the relationship between the equations in the system through the stochastic disturbances analogous to a Seemingly Unrelated Regression (SUR)

# Results

Table 2. Estimates of system of spatial simultaneous equations

	Non-spatial model		Spatial models							
			SLM		SEM		SAC		SDM	
	Housing informality (1)	Labor informality (2)	Housing Informality (3)	Labor informality (4)	Housing informality (5)	Labor Informality (6)	Housing informality (7)	Labor informality (8)	Housing informality (9)	Labor informality (10)
A. Estimates										
Housing informality		1.009*** (0.1171)		1.021*** (0.1178)		0.788*** (0.1223)		0.987*** (0.1200)		1.011*** (0.1315)
Labor informality	0.345*** (0.0408)		0.339*** (0.0415)		0.251*** (0.0420)		0.300*** (0.0400)		0.382*** (0.0517)	
<i>W</i> Housing informality		-0.125 (0.1798)				-0.065 (0.1629)		0.223 (0.7538)	-0.434 (0.3891)	
<i>W</i> Labor informality			0.118 (0.1191)				0.110 (0.1132)	-0.140 (0.1518)	0.465 (0.5339)	
$\rho$				-0.468*** (0.1458)	-0.069 (0.1499)	-0.500** (0.2194)	-0.174 (0.1650)			
R <sup>2</sup>	0.329	0.629	0.333	0.628	0.356	0.652	0.343	0.632	0.343	0.649
Breusch-Pagan test		30.47***		30.47***		13.13***		27.62***		35.85***
N		176		176		176		176		176
LM tests of spatial dependence										
LM <sub>SLM</sub>		4.930*								
Robust LM <sub>SLM</sub>		0.573								
LM <sub>SEM</sub>		7.958**								
Robust LM <sub>SEM</sub>		3.601								
LM <sub>SAC</sub>		8.532*								

Notes: All models include the set of control variables mentioned in Table 1. Standard errors in parenthesis. Spatial models are estimated by Generalized Spatial Three-Stage Least Square (GS3SLS). We instrument labor informality and housing informality variables using the Euclidean distance (in km) from the centroids of each analytical region to the central business district (CBD) of Medellin and the level of housing informality lagged in time (2011), respectively. The inclusion of spatial lag in the models is instrument using the spatial lag of order 2 of control variables. *W* is a matrix (*r* x *r*) of spatial connectivity among analytical regions, which, in our case, is a standardized first-order Queen type.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

- The results for all models consistently indicate that **labor informality has a positive and statistically significant effect on informal housing**, suggesting that precarious labor conditions are a crucial driver of precarious housing conditions
- The estimates reveal that informal housing positively and significantly affects labor informality

# Results

Table 3. Direct and indirect spillovers effects

	SLM			SAC		
	DE	IE	TE	DE	IE	TE
Housing informality						
Labor informality	0.306*** (0.0425)	-0.035 (0.0521)	0.270*** (0.0671)	0.263** (0.0404)	-0.015 (0.0447)	0.247** (0.0638)
Labor informality						
Housing informality	0.895*** (0.1232)	0.134 (0.1457)	1.029*** (0.2190)	0.846*** (0.1242)	0.115 (0.1339)	0.961*** (0.2094)

Notes: DE: Direct effects; IE: Indirect effects; TE: Total effects. Standard errors in parenthesis.

\*\*\*p<0.01. \*\*p<0.05. \*p<0.1

- The estimated direct effects of labor informality on informal housing confirm our previous findings that [informal work activities are critical to the emergence of informal settlements](#)
- The indirect effect (or spillover effect) of labor informality on informal housing is not statistically significant, indicating that the level of informal housing in one region is not affected by the level of labor informality in neighboring regions
- This result confirms the highly localized relationship between informal work activities and informal settlements, implying that many poor workers may find shelter and work opportunities within the slum economy

# Conclusions

- Urban informality is a persistent phenomenon in developing countries and it is an important factor affecting the productivity and wellbeing of cities
- In this study we analyze the relationship between informal housing and labor informality at the intra-urban level in Medellín (Colombia), taking into account the spatial dimension of these two phenomena
- From a point of view of the location of the urban informality, we found:
  - Urban informality is located in poor peripheral areas
  - There is a spatial simultaneity between housing informality and labor informality
  - Housing informality and labor informality are phenomena with a high spatial persistence
- The estimates of the spatial simultaneous equation models showed that labor informality positively affects housing informality  $\implies$  precarious labor conditions are important factors that explain the choice of precarious housing
- Additionally, there are not spatial spillover effects of labor informality on informal housing  $\implies$  This may suggest a highly localized relationship between informal housing and labor informality due to the peculiarities of the informal work carried out within homes, leading to precarious housing conditions

# Housing informality and labor informality in space

## In search of the missing links

Slides in [html](#)

Slides in [PDF](#)

