

Business Visibility and the Measurement of Informality: A Multidimensional Latent Variable Index

Henry Sebastián Rangel-Quiñonez

Parthenope University of Naples

Salvatore Capasso

`salvatore.capasso@uniparthenope.it`

Parthenope University of Naples

Luisa Fernanda Arenas-Estevez

Parthenope University of Naples

Research Article

Keywords: Business informality, Non-concealment, Underreporting and social desirability bias, Latent variable measurement, Confirmatory factor analysis (CFA), Business visibility

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10619388/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Measuring business informality is challenging because many indicators rely on sensitive self-reported information on tax, labour, and registration compliance, which may be affected by non-response, fear of disclosure, or social desirability bias. This paper develops and validates a multidimensional latent index of business informality using firm-level data from EMICRON, Colombia's Microbusiness Survey (2019–2024). Building on DANE's Multidimensional Business Informality Index (IMIE), the study provides, to our knowledge, the first empirical validation of its measurement structure using Confirmatory Factor Analysis and introduces a Non-concealment dimension capturing business visibility and traceability. The final model identifies three dimensions: Inputs, Taxes, and Non-concealment. Compared with alternative measures, the validated model produces substantially higher estimates of informality, exceeding DANE's official measure by about 11 percentage points on average and the dichotomous RUT-based measure by approximately 80%. The results suggest that indicators of business visibility and traceability can complement conventional compliance-based measures and improve the statistical approximation of informal economic activity.

JEL codes: O17,C38, C43

1. Introduction

The concept of the informal economy (IE) has been extensively developed in the economics literature and, in general terms, has been approached from three broad analytical dimensions: the conceptual delimitation of the phenomenon, the development of empirical measurement strategies, and the analysis of its theoretical and empirical relationship with various economic phenomena. The present study contributes to the second of these dimensions by developing and validating a multidimensional instrument to measure business informality.

Strategies for measuring informality can be grouped into two main approaches. The first relies on indirect methods, based on multivariate statistical techniques applied to aggregate data, in which informality is estimated as a latent variable constructed from macroeconomic indicators or from discrepancies between theoretical and observed economic values. Among the most widely used approaches are DYMIMIC and MIMIC models, as well as dynamic general equilibrium frameworks. However, these methods have been criticized for their overly simplifying assumptions and their limited explanatory capacity to explain the relationship between informality and other economic phenomena. In particular, the simultaneous inclusion of multiple macroeconomic indicators often introduces simultaneity-driven endogeneity when these measures are subsequently used in econometric models.

The second approach consists of direct measurement strategies, typically implemented through surveys of firms or workers, which allow the identification of both business and labor informality. Although these methods provide more detailed and disaggregated information, they tend to be more costly in terms of data collection and may suffer from underreporting associated with self-reporting and the sensitivity of certain questions, particularly those related to tax compliance. Complementary strategies have also

emerged, combining direct and indirect approaches, as well as less simple binary classifications in which productive units are labelled as formal or informal based on a single criterion, such as business registration or the payment of social security contributions for employees. However, binary classifications fail to capture the gradual and multidimensional nature of informality, making them less informative than continuous measures of the phenomenon (Capasso et al., 2022; Dell’Anno, 2022; Elgin and Erturk, 2019; OECD, 2002; Schneider, 2023).

A further challenge in measuring business informality arises from the heterogeneity of conceptual definitions regarding what constitutes an informal firm. The notion of informality originates in the concept of “informal sector,” initially understood as the distinction between self-employment and wage employment, according to the degree of connection between economic activity and the state (Hart, 1973; ILO, 1973). Subsequent definitions expanded this perspective. The ILO (2015), for example, associated informality with the absence of coverage under formal institutional arrangements, linking it particularly to small firms with few permanent employees or single-person enterprises. Feige (2016), later, defined informality as the non-compliance with rules governing production and service provision. The absence of a universally accepted definition reflects the diversity of analytical perspectives across economics and statistical sciences and has led to the coexistence of multiple terms, such as the shadow, unofficial, hidden, underground, or black economy (Dell’Anno, 2022). This conceptual diversity is further complicated by regulatory heterogeneity across countries, since the same firm may be classified differently depending on the prevailing legal framework, which complicates the construction of internationally comparable measurement instruments (IMF, 2021; OECD, 2002).

In this context, the validation of measurement instruments becomes particularly important. Statistical validation allows researchers to identify sets of indicators that not only align with theoretical foundations but also exhibit adequate joint statistical properties. This process helps identify latent structures underlying the phenomenon of informality and prevents the inclusion of indicators with weak correlations or poorly justified theoretical relationships (Kalkbrenner, 2021; Yusoff et al., 2021). Moreover, validated instruments improve the reliability of empirical analyses by providing more precise measures of business informality, thereby strengthening the causal models used to study the determinants and consequences of the informal economy. Consequently, the validation of measurement structures represents a crucial step toward constructing robust and adaptable instruments capable of capturing business informality across different institutional contexts.

Despite the growing empirical literature on informality, formal validation exercises for instruments specifically designed to measure business informality remain scarce. Existing studies typically employ indicators that are theoretically associated with informality but do not undergo explicit validation procedures. Ulyssea (2018), for example, distinguishes between an “extensive margin,” defined as the absence of registration and entry fees, and an “intensive margin,” referring to registered firms that keep workers off payroll in Brazil. Hernando Gutierrez and Rodriguez-Lesmes (2023) extend this approach in the Colombian context by classifying firms into three categories – formal, informal extensive, and informal intensive – based on six indicators related to business registration and labor

conditions. Hamaguchi et al. (2025) construct a multidimensional indicator of business informality for Nigeria, Kenya, and Ghana using multiple correspondence analysis applied to six variables related to accounting practices, regulatory compliance, tax payments, business location, acceptance of electronic payments, and registration status. Similarly, Cortés et al. (2025) analyze business informality in Colombia using the Multidimensional Business Informality Index (IMIE) and examine its relationship with variables such as minimum wages and tax rates. However, none of these studies formally validates the measurement structure of the instrument used.

Against this background, the objective of this paper is to develop and empirically validate a multidimensional instrument for measuring business informality in Colombia using firm-level survey data. Building on the Multidimensional Business Informality Index (IMIE) proposed by the National Administrative Department of Statistics (DANE), the study assesses the empirical identifiability of its original factor structure and proposes a refined specification using Confirmatory Factor Analysis. In addition, the paper introduces a new dimension capturing Non-concealment behaviours, which reflect the degree of visibility and traceability of business activity.

The resulting Final validated Model is then compared with four alternative measures of business informality: our own calculation based on the original IMIE structure, an informality proxy derived exclusively from the Non-concealment factor, a simple dichotomous measure based on possession of the Tax Identification Number, RUT, and the official IMIE figures reported by DANE. This comparative exercise is central to the empirical contribution of the paper, as it shows that different measurement choices lead to substantially different estimates of business informality. In particular, the Final validated Model yields consistently higher informality rates than our own calculation based on the original IMIE structure, the official DANE estimates, and the RUT-based measure, suggesting that more restricted approaches may underestimate the extent of business informality.

The contribution of this study is fourfold. First, it provides a formal statistical validation of an instrument specifically designed to measure business informality. Second, it identifies empirical limitations in the original measurement structure and proposes a parsimonious and empirically consistent specification. Third, it expands the measurement framework by introducing indicators associated with business visibility and traceability, which may help mitigate undermeasurement arising from sensitive survey items, high non-response, or social desirability bias. Fourth, it demonstrates empirically that the choice of measurement strategy matters: narrow indicators, such as tax registration alone, capture only part of the phenomenon, whereas a multidimensional latent-variable approach provides a broader and more nuanced representation of business informality. In doing so, the study contributes to the broader literature on the measurement of informality and provides a methodological framework that can be adapted to other countries while preserving conceptual comparability across key dimensions of business formality, including labour obligations, tax compliance, and business visibility or traceability.

The remainder of the paper is organized as follows. Section 2 discusses the conceptual framework and the main approaches to measuring business informality. Section 3 describes the data and the empirical

strategy. Section 4 presents the measurement model and validation strategy. Section 5 reports the results, including the validation of the measurement structure and the comparison between alternative informality measures. Section 6 concludes.

2. Conceptual framework and measurement of business informality

In Colombia, the official measurement of business informality has predominantly relied on direct measurement approaches targeting entrepreneurs. One of the first instruments designed for this purpose was the Micro-establishments Survey, conceived as a follow-up panel and implemented between 2012 and 2016. Its discontinuation was largely due to the fact that, toward the end of the implementation period, most of the observed units corresponded to long-standing businesses, which significantly reduced its capacity to represent the universe of microenterprises (Fernández, 2020).

Since 2019, the Microbusiness Survey (EMICRON) has served as Colombia's main official tool for measuring business informality. Based on a sampling frame derived from the Great Integrated Household Survey (GEIH), EMICRON collects annual information on entrepreneurship, financial inclusion, ICT use, social capital, and informality among microbusinesses. Using this information, DANE constructs the Multidimensional Business Informality Index (IMIE), aligned with CONPES 3956 (Consejo Nacional de Política Económica y Social, 2019).

The conceptual framework underlying the construction of the IMIE is grounded in the institutionalist approach, which interprets informality as the result of transaction costs associated with complying with regulations and standards in lawful economic activities (DANE, 2024). This perspective is closely related to the legalist view, according to which informality emerges as a response to high tax burdens and inefficient or corrupt institutions. In such contexts, entrepreneurs may disengage from formal institutions and conceal part of their economic activity from the State (Dell'Anno, 2022). More broadly, Williams (2008) argues that informality arises from a discrepancy between the norms, values, and beliefs embedded in formal institutions and those prevailing in informal ones.

According to CONPES 3956 (2019), business informality should be understood as a process rather than a binary state, reflecting progressive compliance with applicable regulations. The framework distinguishes four dimensions aligned with the firm's value-generation stages: entry, inputs, production, and taxes. As illustrated in Illustration 1, these dimensions are operationalized through ten indicators designed to capture different aspects of business formalization.

A key conceptual element of this framework is the distinction between informality and illegality. Informality refers to lawful productive activities that fail to comply with certain regulations, whereas illegality involves criminal or expressly prohibited activities. Accordingly, the institutional treatment of informal firms differs from that applied to illicit activities (Corte Constitucional de Colombia, 2013). CONPES 3956 (2019) also specifies that business informality measurement should exclude subsistence economic units, such as own-account workers without paid employees, since they are not considered firms within this framework.

2.1 Measurement of business informality in Latin America

In Latin America, the measurement of business informality is often closely linked to the definition of the informal sector as the set of informal productive units, in line with the guidelines established by the 21st International Conference of Labour Statisticians (ICLS, 2023). In that conference, the International Labour Organization (ILO) defines informal units as those that “are not formally recognized by government authorities as distinct market producers and thus not covered by formal arrangements”. Accordingly, many countries in the region classify informal firms using criteria related to tax registration, legal incorporation, or other forms of official recognition by the State.

As a consequence, several countries have conceptually and statistically separated the different stages of business formality. In many cases, the informality of productive factors, particularly employment, is analyzed independently from the formality status of the productive unit. Under this framework, it is possible for a firm to be classified as formal even when its workers are not. Across the countries reviewed, only Mexico explicitly incorporates employment formality into the firm’s formality classification, integrating business and labor criteria within a single measure (see Annex 1).

In this context, Colombia stands out for having one of the most comprehensive official measures of business informality in the region. Its framework integrates entry requirements, production factors, production processes, and tax compliance within a single multidimensional conceptual structure. Annex 1 summarizes the official criteria and surveys used to measure business and labor informality in major Latin American economies, providing a benchmark that highlights the multidimensional nature of Colombia’s approach (Acuña and Bravo, 2025; Área de Empleo and Universidad de Buenos Aires, 2025; INEGI, 2024; INEI, 2020; Instituto Brasileiro de Geografia e Estatística, 2005; Instituto Nacional de Estadística y Censos, 2015).

3. Data and empirical strategy

The data used in this study come from the EMICRON Survey, conducted by DANE for the period 2019–2024. Based on a sampling frame derived from the GEIH, EMICRON applies a rotation mechanism that excludes previously surveyed segments for three years, reducing repeated observations in nearby periods survey (DANE, 2023a). The consolidated database comprises 485,612 responses from microenterprises with fewer than ten employees. Of these, 46,066 units are classified as employer firms, defined as businesses with at least one paid worker. Since this study focuses on business informality rather than subsistence economic activities, own-account units are excluded from the instrument validation process (DANE, 2022)

Survey responses were provided by the owner of the microbusiness or, when applicable, by one of the partners. Data collection was carried out with the support of DANE officials, through face-to-face interviews in rural areas and telephone interviews in urban areas, ensuring broad territorial coverage.

Table 1 reports the main characteristics of the analytical sample. Most respondents are male, with an average age of 48 years. Most firms operate in urban areas and are concentrated in services and commerce. On average, firms are 14 years old and employ fewer than two paid workers, confirming the small-scale nature of the employer microenterprises analyzed.

Table 1
Sample characteristics (owners and firms)

Characteristics	Mean (SD)	%
Characteristics of the responding owner or partner		
Sex of the owner, male (P35)		74%
Age (P241)	48.1 (13.1)	
Firm characteristics		
Firm age in months (P3034)	168–14 years (139)	
Urban area (Clase)		83%
Industry according to ISIC (GRUPOS4)		
Services		42%
Commerce		23%
Agriculture, livestock, hunting, forestry, and fishing		18%
Manufacturing		14%
Human capital composition		
Workers who receive pay (P3032_1)	1.85 (1.29)	
Partners (P3032_2)	0.10 (0.35)	
Unpaid workers or family workers (P3032_3)	0.113(0.385)	

4. Measurement model and validation strategy

4.1 The IMIE baseline model

The starting point for the construction and validation of the business formality construct is the Multidimensional Business Informality Index (IMIE) developed by DANE for Colombia. Illustration 1 presents the conceptual structure of this index, which comprises four latent factors intended to capture different dimensions of firm informality. Each factor corresponds to a specific stage of business activity and reflects the set of obligations that productive units have toward the Colombian State (DANE, 2024). Although this framework was developed for the Colombian context, its structure can be adapted to other

national settings depending on the regulatory and institutional requirements that define business formalization in each country.

From a theoretical perspective, the model assumes positive relationships among all factors composing the IMIE, as well as between each factor and its corresponding indicators. Higher levels of compliance in one dimension of business formality are therefore expected to be associated with higher levels of compliance in the remaining dimensions. This reflects the idea that formalization (and inversely informality) is a gradual and cumulative process, in which firms progressively comply with different regulatory requirements.

Annex 2 presents the items selected from the EMICRON survey and used for the empirical validation of the IMIE factor structure. In particular, the Input factor, originally represented by two indicators in the conceptual proposal, was operationalized in greater detail by distinguishing labor conditions for workers and owners. This refinement led to the construction of four indicators instead of two, incorporating specific measures of health and pension contributions as well as affiliation with the occupational risks system for both groups. As a result, the original IMIE structure of ten indicators was expanded to twelve indicators for the validation process.

4.2 Extension: IMIE Plus and Non-concealment

Building on the baseline IMIE structure, we propose an extension of the model, referred to IMIE Plus. This extension incorporates additional indicators and introduces a new factor in order to capture dimensions of business informality that are not explicitly considered in the original measurement framework.

First, additional variables are incorporated into the existing factors. In particular, indicators related to social benefits and parafiscal contributions are included in the Input factor, since they represent formal obligations that firms have toward their workers. Furthermore, a variable capturing the tax regime is added to the Taxes factor, given that in some countries, such as Peru, this criterion constitutes a key element in defining business formality (INEI, 2020).

In addition to these adjustments, the model introduces a new latent factor termed Non-concealment, which expands the operational definition of business informality. While the original IMIE framework focuses on compliance with formal legal requirements imposed on firms in Colombia, the Non-concealment factor aims to capture behaviors that intentionally reduce the visibility of economic activity and hinder state oversight. The literature has emphasized that such practices are closely associated with informality because they limit the traceability of business operations and complicate monitoring by the authorities (Capasso and Jappelli, 2013; Dell’Anno, 2022; Medina and Schneider, 2018).

From this perspective, concealment behaviors may appear conceptually inconsistent with the expected conduct of formal firms, which typically seek to strengthen their visibility and reputation in the market. However, these behaviors can be understood as rational responses to regulatory or tax burdens, particularly in contexts where the perceived costs of formalization are high. In practice, concealment may manifest through several observable strategies, including the exclusive use of cash as a means of

payment, the absence of advertising or professional services, the lack of a website or trade name, operating from private or non-commercial locations, functioning as a street vendor, or maintaining limited public visibility. All these practices reduce the likelihood that a firm will be identified and monitored by the competent authorities.

Firm size is also included among the indicators of this factor, measured through the number of workers employed. In several countries, including Brazil and Mexico, firm size is used as a proxy for business formality (Área de Empleo and Universidad de Buenos Aires, 2025; Instituto Brasileiro de Geografia e Estatística, 2005). In this study, this variable is incorporated into the Non-concealment factor, because the scale of the productive unit directly influences the ability of firms to conceal their activities. Smaller firms generally have greater opportunities to operate outside the institutional radar, whereas increases in employment tend to raise visibility and reduce the scope for concealment.

The factorial proposal corresponding to the IMIE Plus, Non-concealment model is presented in Illustration 2, while Annex 3 reports the EMICRON survey items used for its construction. By integrating both regulatory compliance indicators and behavioral proxies related to business visibility, this extended framework provides a more comprehensive representation of the multidimensional nature of business informality.

4.3 Confirmatory factor analysis and validation criteria

Confirmatory Factor Analysis (CFA) was used to validate the proposed measurement models. To assess the empirical feasibility of the factor structures and reduce the likelihood of improper solutions associated with highly correlated indicators, the correlation matrix was examined prior to model estimation in order to detect potential redundancy, collinearity, or lack of common variance among indicators (Brown, 2015). The validation process began with the estimation of the original IMIE structure proposed by DANE. Once the model with the best empirical fit within this baseline framework was identified, the new items and the Non-concealment factor were incorporated progressively, giving rise to the expanded versions of the index.

Model estimation was conducted using the lavaan package in R software (version 4.2.1) (Rosseel, 2012). Given the categorical and ordinal nature of the observed variables, the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator was employed. This estimator is based on polychoric correlations and provides robust standard errors, as well as fit statistics corrected for non-normality (Rosseel, 2025). For identification purposes, the variances of the latent factors were fixed at one, allowing the factor loadings of the associated indicators to be freely estimated.

Following the literature, factor loadings above 0.5 were considered acceptable in order to ensure robust measurement of the latent constructs Yusoff et al. (2021), a threshold stricter than the minimum value of 0.3 sometimes adopted in applied research (Brown, 2015).

Model fit was evaluated using absolute (RMSEA and SRMR), incremental (CFI, TLI, and GFI), and comparative (ECVI) fit indices, following conventional thresholds reported in the literature (Bentler, 1990; Bentler and Bonett, 1980; Brown, 2015; Schermelleh-Engel et al., 2003). The chi-square statistic (χ^2) was also reported for reference, although its interpretation is sensitive to sample size (Brown, 2015; Pituch and Stevens, 2015). When appropriate, scaled indices were reported to account for the categorical nature of the indicators.

The instrument's internal reliability was evaluated using two complementary coefficients. Cronbach's alpha (α) was computed as a conventional indicator of internal consistency, while McDonald's omega (ω) was also reported to account for the limitations of alpha when applied to categorical or ordinal data (Revelle, 2022). Following common practice in the literature, values of 0.7 or higher were considered acceptable for both coefficients (Brown, 2015; Kalkbrenner, 2021) -

Finally, temporal stability was examined through multigroup factorial invariance tests by year, estimated using the semTools package in R (Jorgensen et al., 2025). A hierarchical sequence of nested models, including configural, metric, scalar, and means invariance, was evaluated. Model comparisons focused primarily on changes in CFI ($\Delta\text{CFI} \leq 0.01$), and invariance models were estimated using the WLSMV estimator, reporting both scaled and absolute fit indices to ensure robust intertemporal comparisons.

4.4 Comparison of alternative informality measures

After validating the measurement structure, the study compares the formality scores derived from the Final validated Model with alternative measures of business informality. This comparison is intended to assess how different operational definitions of informality lead to different empirical classifications of firms. The formality score derived from the Final validated Model is computed using an equal-weighting strategy, both across items within each factor and across the three factors. This means that each factor contributes equally to the final score, and each item contributes equally within its corresponding factor. This decision was adopted to facilitate intertemporal comparability and to avoid assigning empirical weights that may vary across years.

Four reference measures are considered. First, we compute our own calculation based on the original IMIE structure, following the conceptual dimensions and weighting logic proposed by DANE (2025). Second, we derive an informality proxy from the Non-concealment factor, using only indicators related to the visibility and traceability of business activity. Third, we construct a simple dichotomous measure based on possession of the Tax Identification Number, RUT, which represents an entry-based criterion of formality. Finally, we compare these estimates with the official IMIE figures reported by DANE for employer firms (DANE, 2023b, 2025). For the continuous measures, firms are classified as informal using a common threshold of 0.5 applied to the estimated formality scores. Specifically, firms with formality scores below 0.5 are classified as informal, while firms with scores equal to or above 0.5 are classified as formal. All estimates are adjusted using the national expansion factor in order to obtain population-level informality rates.

5. Results

5.1 Descriptive statistics

Table 2 reports the descriptive statistics for the dichotomous indicators used in the construction and validation of the instrument. Overall, the variables associated with the Entry factor show moderate compliance rates. In particular, 60.2% of firms report having a Tax Identification Number, RUT (X1), while only 42.3% report being registered with a Chamber of Commerce (X2). This difference is noteworthy given that obtaining the RUT is a free procedure in Colombia, whereas the cost of commercial registration ranges between 10.45 and 20.86 euros for firms with assets below 17,000 euros in 2025¹ (Confecámaras, 2024). Compared with the statutory daily minimum wage in Colombia for the same year (10.76 euros), these costs remain relatively low, suggesting that aversion to formalization persists even when economic entry barriers are limited.

Table Figure 2 **Descriptive statistics of dichotomous indicators**

Name	% Yes (1)	% NA
X1 Tax Identification Number (RUT) P1633	60.2	0
X2 Commercial registry, P1055	42.3	0
X3 Health and pension contributions (workers), P3080_score	18	0.004
X4 Occupational risk insurance (workers), P3084_score	16.2	0.004
X4b Occupational risk insurance (owner), P3090	17.8	0
X5 Renewal of the commercial registry, P661	32.3	0
X6 Operational licenses and fees, P3017_I + P3017L	30.2	0
X11 Social benefits, P3082_score	9.57	0.007
X12 Parafiscal contributions, P2990_score	9.62	11.8
X13 Tax regime, P986	15.7	39.8
X14 Has trade name, P3035	49.6	0
X16 Visible to the public, P469	76.3	17.2
X17 Website, online presence, P2532	6.33	0
X18 Advertising or professional services, P2532	10.1	0
X19 Number of workers greater than 5, (P3032_1 + P3032_3)	3.1	0

The variables corresponding to the Input factor exhibit particularly low compliance levels. The indicators related to workers’ health and pension contributions (X3), as well as affiliation with the occupational risk insurance system for workers (X4) and owners (X4b), all show percentages below 18%, indicating limited

compliance with basic labour obligations even among firms employing paid workers. Indicators associated with the Non-concealment factor display more heterogeneous patterns, ranging from 76.3% of firms reporting that their activity is visible to the public (X16) to only 6.3% reporting an online presence (X17), suggesting important differences in the visibility and traceability of business activity. Table 3 reports the distribution of the polytomous variables, which capture different levels of compliance with regulatory obligations and business practices associated with formality. For indicator X3b, corresponding to the owner's payment of health and pension contributions, only 24% of entrepreneurs report paying both contributions, while 7.37% report paying at least one of them.

Table 3 Distribution of polytomous variables

Categorical variables	0%	1%	2%	NA %
X3b Health and pension contributions (owner) P3088	68.6	7.37	24	0
X7 Industry and Commerce Tax (ICA) P2991	10.3	19.7	20.7	49.3
X8 Value-Added Tax (VAT) P2992	1.17	44.8	6.55	47.5
X9 Income tax return, P2993	10.2	22.3	18.2	49.3
X15 Main location, P3053	15.4	41.8	42.9	1.19
X10 Accounting and financial records, P640	32.7	42.9	24.4	0
X20 Non-cash payment methods, P1764_1	47	23.2	7.84	22

Tax-related indicators reveal substantial heterogeneity. While non-compliance levels for the ICA and income tax remain below 20%, a large proportion of firms report not being required to file certain taxes, particularly VAT. The results also show limited use of formal accounting practices and a strong dependence on cash transactions. Only 7.84% of firms accept a broad range of payment instruments, whereas 47.0% operate exclusively with cash.

It should be noted that X20 was not measured in 2020, as the survey temporarily replaced several questions with a module on firms' adaptive expectations during the COVID-19 pandemic. Consequently, when models include this variable, observations for 2020 are excluded; otherwise, they are retained in the validation process.

5.2 Diagnosis of the IMIE structure

The tetrachoric and polychoric correlation matrix of the IMIE indicators is reported in Annex 4. The analysis reveals several patterns that affect the empirical identifiability of the original measurement structure. First, the indicators associated with the Entry factor exhibit very high correlations. In particular, the Tax Identification Number (X1) and Chamber of Commerce registration (X2) show a correlation of approximately 0.93, while X2 is almost perfectly correlated with renewal of the commercial registry (X5), with a coefficient close to 0.99. These results suggest substantial redundancy among indicators

intended to capture business registration and legal recognition. Second, the Tax Identification Number (X1) displays virtually no correlation with the tax compliance indicators (X7–X9). This pattern is largely explained by the survey design, as firms without a Tax Identification Number were not asked the tax-related questions. Consequently, the Entry and Taxes dimensions do not share sufficient common variance to support a coherent latent structure. A similar issue is observed within the Inputs factor. The indicator measuring workers’ affiliation with the occupational risk insurance system (X4) is almost perfectly correlated with workers’ health and pension contributions (X3), indicating that both variables capture nearly identical information.

Taken together, these findings point to two main empirical problems: redundancy among several indicators and insufficient shared variance between some dimensions. To improve model identifiability and parsimony, X1 and X2 were excluded from the Confirmatory Factor Analysis. Because removing X1 would leave the Entry factor represented by a single indicator, the entire factor was discarded from the validation process. In addition, X4 was removed due to its near-perfect correlation with X3. These adjustments preserve the conceptual coherence of the instrument while reducing collinearity and improving the empirical feasibility of the measurement model.

5.3 Improved IMIE model

The original IMIE structure follows the theoretical proposal formulated by DANE (2024), with the exclusion of the Entry factor due to the convergence issues identified during estimation and documented in the preliminary diagnosis. The adjusted specification, referred to as Improved IMIE, retains X3 and X3b in the Inputs factor, X5 and X6 in the Product factor, and X7–X10 in the Taxes factor.

The Improved IMIE version introduces minor adjustments aimed at improving model identifiability and parsimony while preserving the conceptual structure of the index. In particular, X4b (owner affiliation with the occupational risks system) is excluded due to its high correlation with X3b. The latter is retained because, together with X3, it more directly captures labour conditions associated with formal employment. Although X6 shows a relatively high correlation with X5, both indicators are retained within the Product factor, as removing either variable would leave the factor under-identified at this stage of the analysis.

Standardized factor loadings estimated for both specifications are presented in Table 4. All indicators exhibit loadings above the 0.5 threshold, indicating satisfactory relationships between the observed variables and their corresponding latent constructs. The lowest loading corresponds to X9 (income tax return) within the Taxes factor, although it remains within acceptable ranges.

Table 4 **Standardized factor loadings for the original IMIE and Improved IMIE**

Factor	Indicator	Original IMIE λ	Improved IMIE λ
Inputs	X3	0.825***	0.903***
	X3b	0.955***	0.782***
	X4b	0.923***	—
Product	X5	0.978***	0.976***
	X6	0.901***	0.903***
Taxes	X7	0.771***	0.765***
	X8	0.718***	0.727***
	X9	0.653***	0.652***
	X10	0.715***	0.715***

*** $p < 0.001$.

Overall model fit statistics are reported in Table 5. Relative to the original specification, the Improved IMIE model shows consistent improvements across several fit indices. In particular, RMSEA, SRMR, and ECVI decrease, indicating a better fit of the adjusted structure, while incremental fit indices remain above conventional thresholds.

Table5 **Model fit indices for the original IMIE and Improved IMIE**

Index	Original IMIE	Improved IMIE
χ^2	3786.8	2323.3
χ^2 (scaled)	5405.1	3671.6
df	24	17
CFI	0.986	0.987
TLI	0.980	0.979
CFI (scaled)	0.971	0.972
TLI (scaled)	0.956	0.954
RMSEA	0.083	0.077
RMSEA (scaled)	0.099	0.097
SRMR	0.069	0.060
GFI	0.989	0.990
ECVI	0.169	0.104
N	22,742	22,742

The lower ECVI value observed in the Improved IMIE specification further supports its superior generalization capacity relative to the original model. Correlations among the latent factors, reported in Annex 5, remain below the 0.80 threshold, suggesting that the factors capture distinct dimensions of business formality. Nevertheless, their magnitude indicates the possible presence of a common second-order factor, interpretable as a general formality construct, which is explored in subsequent model extensions.

Finally, the internal reliability of the Improved IMIE factors is generally acceptable. The Inputs–Product block exhibits high reliability ($\alpha = 0.85$, $\omega = 0.85$), while the second component of this block shows values of $\alpha = 0.70$ and $\omega = 0.70$. In contrast, the Taxes factor presents slightly lower reliability ($\alpha = 0.68$, $\omega = 0.69$), falling marginally below the conventional threshold.

5.4 Extended model: IMIE Plus Non-concealment

The IMIE Plus Non-concealment specification extends the Improved IMIE model by incorporating the Non-concealment factor, together with the additional indicators X11 (social benefits), X12 (para-fiscal contributions) and X13 (tax regime), as proposed in Illustration 2. This extension seeks to complement the measurement of business informality by incorporating behaviors associated with the deliberate concealment of economic activity, which may facilitate the evasion of regulatory and fiscal obligations.

However, the preliminary analysis of the tetrachoric and polychoric correlation matrix (see Annex 5) reveals very high correlations among X3, X11, and X12, indicating substantial redundancy of information. Consequently, X11 and X12 were excluded from the validation process, while X3 was retained to preserve coherence with the original IMIE structure and avoid collinearity problems. After this refinement, the model comprised four factors. The Inputs factor included X3 and X3b; the Product factor included X5 and X6; the Taxes factor included X7, X8, X9, X10, and X13; and the Non-concealment factor included X14, X15, X16, X17, X18, X19, and X20.

Model estimation shows that most indicators display statistically significant factor loadings of moderate to high magnitude. However, two variables within the Non-concealment factor, X16 (visible to the public) and X19 (number of workers greater than five), exhibit relatively low loadings, suggesting limited shared variance with the latent construct.

Additionally, the Product factor shows high correlations between its indicators. Given that **X5 (renewal of the commercial registry)** demonstrates stronger empirical performance in terms of factor loadings and model stability, **X6 was excluded** to avoid redundancy.

The correlations among latent factors (Annex 6) also reveal a strong association between the Product and Non-concealment factors (0.727), suggesting that both capture related aspects of firms' interactions with regulatory authorities. In this context, reallocating X5 to the Non-concealment factor is consistent both empirically and conceptually, since obtaining or renewing the commercial registry increases the visibility and traceability of firms.

Standardized factor loadings for the IMIE Plus Non-concealment model are presented Table 6. The Inputs and Taxes factors show relatively high loadings, while the Non-concealment factor shows moderate loadings, reflecting the heterogeneous nature of concealment behaviors.

Table 6 **Standardized factor loadings for IMIE Plus Non-concealment**

IMIE Plus Non-concealment		
Factor	Indicator	Standardized loading (λ)
Inputs	X3	0.908***
	X3b	0.771***
Product	X5	0.969***
	X6	0.900***
Non-concealment	X14	0.625***
	X15	0.532***
	X16	0.412***
	X17	0.536***
	X18	0.703***
	X20	0.552***
	X19	0.401***
Taxes	X7	0.746***
	X8	0.661***
	X9	0.674***
	X10	0.732***
	X13	0.823***

*** $p < 0.001$.

Based on these results, a further refinement of the model is proposed, consisting of merging the Product and Non-concealment factors and removing indicators X16 and X19, which contribute limited explanatory power to the measurement structure.

5.5 Final validated Model

Illustration 3 presents the Final Model, derived from the refinements introduced in the IMIE plus Non-concealment specification. In this solution, the Inputs and Taxes factors exhibit relatively strong loadings, ranging from 0.65 and 0.91, while the Non-concealment factor shows moderate loadings between 0.54 and 0.77.

Correlations among the latent factors are moderate: 0.59 between Non-concealment and Inputs, 0.72 between Taxes and Inputs, and 0.84 between Non-concealment and Taxes. These associations suggest

the presence of a common latent dimension underlying the three components, which may be interpreted as a broader construct of business informality.

The thresholds parameters reported in Illustration 3 provide additional insights into the behavior of certain indicators. In the case of X3b, the proximity of the estimated thresholds (0.11 and 0.365) indicates that the variable effectively operates with two meaningful categories: firms that do not pay health or pension contributions and those that pay at least one of them. Conversely, the strongly negative threshold observed for X8 (-2.035) suggests that firms with very low levels of formality tend to transition from the category “does not pay” to “not required to pay.” This pattern supports merging the first two categories into a single category representing firms that do not pay or are not required to pay VAT, versus those that effectively pay the tax.

Model fit statistics for the IMIE plus Non-concealment and the Final Model are reported in Table 7. Overall, the Final Model shows improvements in several fit indices, including increases in CFI, GFI and TLI, as well as reductions in SRMR, and ECVI, indicating a better overall fit. Although RMSEA, both scaled and unscaled, does not improve in the same direction, its values remain within the acceptable range for the model evaluation.

Table 7 Fit indices for IMIE plus Non-concealment and the Final Model

Index	IMIE Plus Non-concealment	Final Model
χ^2	8,430.1	6,372.4
χ^2 (scaled)	10,565.6	8,495.8
df	98	62
p-value (χ^2)	< 0.001	< 0.001
CFI	0.956	0.963
TLI	0.946	0.954
CFI (scaled)	0.917	0.929
TLI (scaled)	0.899	0.910
RMSEA	0.076	0.078
RMSEA (scaled)	0.085	0.090
SRMR	0.101	0.087
GFI	0.971	0.976
ECVI	0.574	0.382
N	14837	16857

Internal consistency measures indicate that the Final Model achieves acceptable reliability across its three dimensions. The Input factors shows reliability values of α : 0.7 and ω : 0.7, the Non-concealment factor α : 0.7 and ω : 0.71, and the Taxes factor α : 0.75 and ω : 0.75, meeting the conventional 0.7 threshold for internal consistency.

Finally, the results of the **factorial invariance tests by year** (Annex 7). indicate that the model maintains acceptable temporal stability, particularly when evaluated using scaled indices. This finding suggests that the expanded IMIE index can be used reliably in intertemporal comparisons, both in terms of structural relationships and latent means.

5.6 Comparison of Formality Measures

Illustration 4 presents the annual density distributions of the formality scores derived from the Final validated Model, the original IMIE structure, and the Non-concealment factor. The specific weights used to construct the Final validated Model are reported in Annex 8. Across all years, the three measures show a high concentration of firms at low levels of formality, indicating that a large share of employer microenterprises comply with only a limited set of formal requirements. Although the distributions vary moderately over time, their overall shapes remain relatively stable, suggesting persistent patterns of business informality between 2019 and 2024.

The density distribution based on the original IMIE structure exhibits more pronounced local peaks and abrupt changes in shape than the other measures, suggesting clustering around specific score values. This pattern is consistent with the empirical issues identified during the validation process, particularly the presence of redundant indicators and weak common variance among some dimensions. By contrast, the Final validated Model displays a smoother distribution with a greater representation of intermediate levels of formality, indicating an improved capacity to capture heterogeneity among employer microenterprises.

The Non-concealment factor exhibits an intermediate distributional pattern and substantial overlap with the original IMIE structure (0.74–0.78 across years). However, unlike the Inputs and Taxes dimensions, which rely primarily on self-reported compliance with labour, accounting, and tax obligations, Non-concealment captures characteristics related to the visibility and traceability of business activity. Because these indicators involve less sensitive information, they may be less affected by non-response, reporting errors, or social desirability bias. Consequently, incorporating this dimension may improve the identification of informal firms when direct compliance measures are incomplete or potentially biased.

Table 8 presents the estimated percentage of business informality under the different measurement strategies. The Final validated Model consistently produces the highest informality estimates, declining from 88.9% in 2019 to 79.9% in 2024, but remaining persistently high throughout the period.

Table 8 Comparison of business informality measures

Year	Informality, Final validated Model (%)	Informality from the Non-concealment factor (%)	Informality from our own calculation based on the original IMIE structure (%)	Official IMIE DANE report (%)	Informality according to RUT (%)
2019	88.9	75.8	81.1	71.7	45.3
2020	82.0	73.8	70.5	70.5	43.7
2021	83.0	72.6	68.9	66.9	40.7
2022	80.9	73.0	68.6	66.3	44.8
2023	78.3	72.1	68.8	67.7	43.8
2024	79.9	72.4	68.6	67.9	46.1
Items	13	6	10	10	1

Note: Values adjusted using the national expansion factor.

The difference between the Final validated Model and our own calculation based on the original IMIE structure is substantial. On average, the validated model estimates informality levels that are 11.08 percentage points higher. By incorporating indicators related to business visibility and traceability, the validated model captures additional manifestations of informality that are less dependent on sensitive compliance declarations.

A relatively small average difference of 1.2 percentage points is observed between the official IMIE figures reported by DANE and our own calculation based on the original IMIE structure for the 2020–2024 period, suggesting that the calculation implemented in this study broadly reproduces the official measurement logic. The larger difference observed in 2019 (9.4 percentage points) may reflect differences in data cleaning, preprocessing decisions, or the treatment of variables associated with the Inputs factor. The comparison with the RUT-based measure highlights the limitations of using a single entry criterion to approximate business informality. Across the period, informality according to the Final validated Model is approximately 1.8 times higher than the estimate obtained from the absence of a Tax Identification Number, indicating that tax registration captures only one dimension of formality.

The Non-concealment factor provides an intermediate estimate between the Final validated Model and the official IMIE figures. Its informality rate ranges from 75.8% in 2019 to 72.4% in 2024 and is, on average, 4.8 percentage points higher than the official DANE estimates. Although it does not capture the full multidimensional structure of the Final validated Model, its stability and proximity to broader measures suggest that it may serve as a useful complementary proxy. Overall, the comparison shows that measurement choices have important implications for the estimated magnitude of business informality. Measures based on isolated criteria, such as tax registration alone, produce substantially lower estimates, whereas multidimensional approaches incorporating labour, tax, accounting, and visibility-related dimensions provide a broader representation of the phenomenon.

6. Conclusions

This study develops and validates an alternative multidimensional instrument to measure business informality in Colombia as a latent and continuous construct, using data from the 2019–2024 Microbusiness Survey. The contribution of the paper goes beyond the statistical validation of the Multidimensional Business. It also proposes a refined measurement framework that corrects empirical limitations in the original structure and incorporates business visibility and traceability as complementary dimensions for approximating informal activity. This is particularly relevant because conventional compliance-based measures may underestimate informality when firms underreport sensitive information on tax obligations, labour contributions, or registration status. The resulting instrument combines regulatory compliance and Non-concealment indicators into a more parsimonious and empirically consistent measure, aligned with the institutionalist approach and the CONPES 3956 framework. In this way, the paper provides both an assessment of an existing official measurement structure and an alternative measure that can be adapted to other institutional contexts. Annex 9 presents the questionnaire corresponding to the validated Final Model.

The empirical analysis shows that the original IMIE structure presents identifiability limitations associated with redundancies and insufficient common variance among several indicators. Addressing these issues through correlation diagnostics and parsimony criteria leads to an adjusted specification, termed Improved IMIE, which exhibits better empirical performance while preserving the conceptual foundations of the index and maintaining satisfactory factor loadings across retained indicators. Building on this refinement, the Final Model identifies a three-factor structure capturing the main dimensions of business informality: Inputs, related to labour formality and social security contributions; Taxes, associated with tax compliance and accounting practices; and Non-concealment, reflecting the degree of visibility and traceability of business activity through observable characteristics such as renewal of the commercial registry, trade name, main business location, public visibility, online presence, advertising or professional services, and the acceptance of non-cash payment methods.

Consistent with the literature on the informal economy and concealment (Capasso and Jappelli, 2013; Dell’Anno, 2022, 2023; Medina and Schneider, 2018), the results indicate that these dimensions are interrelated rather than independent. In particular, the moderate correlations among latent factors, especially between Taxes and Inputs and between Taxes and Non-concealment, suggest the existence of a common underlying component interpretable as a latent construct of informality shared across different domains of compliance and traceability. The comparison of informality measures further shows that the Final validated Model estimates consistently higher levels of business informality than our own calculation based on the original IMIE structure, the official IMIE figures reported by DANE, and the simple dichotomous measure based on possession of the Tax Identification Number, RUT. This finding suggests that estimates based mainly on self-reported compliance questions may be downwardly affected by the sensitivity of items related to tax obligations, labour contributions, and formal registration. In this context, social desirability bias or fear of disclosure may lead to lower estimated levels of informality. At the same time, the comparison with the RUT-based measure confirms

the limitations of using a single entry criterion to approximate business informality. Tax registration captures only one dimension of formality and overlooks relevant aspects related to labour obligations, tax practices, accounting records, business visibility, and traceability. Therefore, the results support the need to measure business informality as a multidimensional phenomenon rather than as a binary condition based exclusively on registration status.

An important applied implication of the findings is that the Non-concealment dimension may serve as a complementary proxy for business informality in contexts where sensitive survey items, such as tax compliance or labour contributions, are prone to underreporting, present high levels of non-response, or are not consistently available across years. Although this factor is also based on self-reported survey information, it captures characteristics related to the visibility and traceability of business activity rather than direct declarations of regulatory compliance. As a result, these indicators may be less affected by social desirability bias or fear of disclosure and can help approximate informality when firms do not respond to the Taxes or Inputs modules, or when compliance-related responses are potentially biased. While these indicators do not replace direct measures of regulatory compliance, they can attenuate some limitations of self-reported data and strengthen the empirical approximation of business informality within the measurement framework proposed in this study.

Beyond the Colombian case, the methodological approach proposed in this study provides a replicable framework for the development and validation of multidimensional measures of business informality. In particular, the procedure offers a basis for adapting the construct to different institutional contexts while preserving conceptually comparable dimensions. At the same time, the specific indicators can be tailored to reflect the regulatory requirements of each national context.

Finally, the interpretation of the results should consider some limitations. Although the Non-concealment factor captures aspects of business visibility and traceability, it should not be regarded as a substitute for direct measures of regulatory compliance when such information is available. Rather, its main value lies in complementing conventional measures of informality and providing an alternative source of information in contexts where data are unavailable, incomplete, or potentially affected by reporting bias. Accordingly, the Non-concealment dimension should be interpreted as a complementary proxy that can strengthen, but not replace, direct measures of business formality. Future research could examine how multidimensional measures of business informality, as well as proxy dimensions such as Non-concealment, are associated with variables widely studied in the literature, including financial development, productivity, technological adoption, firm growth, access to credit, and business performance. This would allow future studies to assess not only the validity of the proposed instrument, but also its explanatory capacity in relation to key economic outcomes. In addition, future research could explore the heterogeneity of firms according to their profiles across the different dimensions of the index. For example, it would be useful to identify firms with low levels of Non-concealment but higher levels of compliance in the remaining dimensions, as well as firms with low levels of visibility and traceability together with low levels of tax or labour compliance. Other combinations of scores across Inputs, Taxes, and Non-concealment could also be used to construct firm typologies. This approach

would make it possible to examine whether different empirical profiles correspond to theoretical interpretations of informality discussed in the literature, such as voluntary non-compliance or free-riding, exclusion- or subsistence-driven informality, and informality associated with institutional mismatch or distrust in formal rules.

Declarations

Author Contribution

H.S.R.Q. led the conceptualization, methodology, investigation, formal analysis, data curation, and preparation of the original manuscript. L.F.A.E. contributed to the conceptualization, preparation of the original manuscript, and review and editing. S.C. contributed to the conceptualization, supervision, and critical review and editing of the manuscript. All authors contributed to the writing of the manuscript, reviewed the final version, and approved it for publication.

Acknowledgement

The authors are grateful to Professor Guglielmo Maria Caporale of Brunel University London for his valuable comments and suggestions on earlier versions of this manuscript.

Data Availability

The data supporting the findings of this study are publicly available from the National Administrative Department of Statistics of Colombia, DANE, through its official microdata repository. The study uses the Microbusiness Survey, EMICRON, for the period 2019–2024. The datasets are available in the public domain at DANE’s microdata catalogue for each survey year: 2019, 2020, 2021, 2022, 2023, and 2024.- Emicron 2024, <https://microdatos.dane.gov.co/index.php/catalog/875>- Emicron 2023, <https://microdatos.dane.gov.co/index.php/catalog/832>- Emicron 2022, <https://microdatos.dane.gov.co/index.php/catalog/796>- Emicron 2021, <https://microdatos.dane.gov.co/index.php/catalog/742>- Emicron 2020, <https://microdatos.dane.gov.co/index.php/catalog/712>- Emicron 2019, <https://microdatos.dane.gov.co/index.php/catalog/681>

References

1. Área de Empleo, distribución e instituciones L. and Universidad de Buenos Aires. (2025), *Panorama Del Empleo Informal y La Pobreza Laboral*.
2. Bentler, P.M. (1990), “Comparative fit indexes in structural models”, *Psychological Bulletin*, American Psychological Association Inc., Vol. 107 No. 2, pp. 238–246, doi: 10.1037/0033-2909.107.2.238.

3. Bentler, P.M. and Bonett, D.G. (1980), "Significance tests and goodness of fit in the analysis of covariance structures", *Psychological Bulletin*, Vol. 88 No. 3, pp. 588–606, doi: 10.1037/0033-2909.88.3.588.
4. Brown, T.A. (2015), *Confirmatory Factor Analysis for Applied Research*, edited by Little, T.D., second., The Guilford Press.
5. Capasso, S. and Jappelli, T. (2013), "Financial development and the underground economy", *Journal of Development Economics*, Vol. 101 No. 1, pp. 167–178, doi: 10.1016/j.jdeveco.2012.10.005.
6. Capasso, S., Ohnsorge, F. and Yu, S. (2022), "Informality and financial development: A literature review", *Manchester School*, John Wiley and Sons Inc, 1 September, doi: 10.1111/manc.12417.
7. Confecámaras. (2024), *Tarifas de Registro Mercantil, Entidades Sin Ánimo de Lucro, Proponentes y RUNEOL Vigencia 2025*.
8. Consejo Nacional de Política Económica y Social. (2019), *CONPES 3956*, Bogotá.
9. Corte Constitucional de Colombia. (2013), *Sentencia T-386 de 2013*.
10. Cortés, A.O., Acosta, K., Arango, L.E., Aristizábal, D., Ávila, Ó., Becerra, Ó., Fernández, C., et al. (2025), "New evidence on labor and business informality in Colombia", *Ensayos Sobre Política Económica*, Banco de la República, Vol. 2025 No. 108, doi: 10.32468/espe108.
11. DANE. (2022), *Manual de Recolección y Conceptos Básico de La Encuesta de Micronegocios*.
12. DANE. (2023a), *METODOLOGÍA GENERAL ENCUESTA DE MICRONEGOCIOS-EMICRON*.
13. DANE. (2023b), *Índice Multidimensional de Informalidad Empresarial (IMIE) 2022*.
14. DANE. (2024), *Índice Multidimensional de Informalidad Empresarial (IMIE) 2023*, Bogotá.
15. DANE. (2025), *Índice Multidimensional de Informalidad Empresarial (IMIE) 2024*.
16. Dell'Anno, R. (2022), "Theories and definitions of the informal economy: A survey", *Journal of Economic Surveys*, John Wiley and Sons Inc, Vol. 36 No. 5, pp. 1610–1643, doi: 10.1111/joes.12487.
17. Dell'Anno, R. (2023), "Measuring the unobservable: estimating informal economy by a structural equation modeling approach", *International Tax and Public Finance*, Vol. 30 No. 1, pp. 247–277, doi: 10.1007/s10797-022-09742-0.
18. Elgin, C. and Erturk, F. (2019), "Informal economies around the world: measures, determinants and consequences", *Eurasian Economic Review*, Springer International Publishing, Vol. 9 No. 2, pp. 221–237, doi: 10.1007/s40822-018-0105-5.
19. Feige, E.L. (2016), "Reflections on the meaning and measurement of unobserved economies: What do we really know about the 'Shadow Economy'?", *Journal of Tax Administration*, Vol. 2 No. 1, pp. 5–41.
20. Fernández, C. (2020), "Informalidad Empresarial en Colombia", *COYUNTURA ECONÓMICA*, Vol. 50, pp. 133–168.
21. Hamaguchi, N., Hino, H., Piot, C. and Yin, J. (2025), "Multi-dimensional informality and heterogeneity of microenterprises in urban Africa", *Japanese Economic Review*, Springer, doi: 10.1007/s42973-025-00212-w.

22. Hart, Keith. (1973), "Informal income opportunities and urban employment in Ghana", *The Journal of Modern African Studies*, Vol. 11, pp. 61–89.
23. Hernando Gutierrez, L. and Rodriguez-Lesmes, P. (2023), "Productivity gaps at formal and informal microfirms", *World Development*, Elsevier Ltd, Vol. 165, doi: 10.1016/j.worlddev.2023.106205.
24. ICLS. (2023), *Resolution Concerning Statistics on the Informal Economy*, Genova.
25. ILO. (1973), *Employment, Incomes and Equality a Strategy for Increasing Productive Employment in Kenya*.
26. ILO. (2015), *National Employment Policies: A Guide for Workers' Organisations*.
27. IMF. (2021), *Measuring the Informal Economy*, Washington, D.C.
28. INEI. (2020), *Producción y Empleo Informal En El Perú. Cuenta Satélite de La Economía Informal 2007-2019*, Lima.
29. Instituto Brasileiro de Geografia e Estatística. (2005), *Economía Informal, Urbana*, IBGE, Rio de Janeiro.
30. Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M. and Rosseel, Y. (2025), "semTools: Useful tools for structural equation modeling".
31. Kalkbrenner, M.T. (2021), "A Practical Guide to Instrument Development and Score Validation in the Social Sciences: The MEASURE Approach", *Practical Assessment, Research, and Evaluation*, University of Massachusetts Amherst Libraries, Vol. 26 No. 1, doi: 10.7275/SVG4-E671.
32. Medina, L. and Schneider, F. (2018), *Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?*
33. OECD. (2002), *Measuring the Non-Observed Economy: A Handbook*, OECD.
34. Pituch, K.A. and Stevens, J. (2015), *Applied Multivariate Statistics for the Social Sciences*, *Applied Multivariate Statistics for the Social Sciences*, doi: 10.4324/9781315814919.
35. Revelle, W. (2022), "psych: Procedures for Psychological, Psychometric, and Personality Research", Evanston, Illinois.
36. Rosseel, Y. (2012), "lavaan: An R Package for Structural Equation Modeling", *Journal of Statistical Software*, Vol. 48 No. 2, pp. 1–36, doi: 10.18637/jss.v048.i02.
37. Rosseel, Y. (2025), "'Lavaan' Latent Variable Analysis", CRAN, 21 September.
38. Schermelleh-Engel, K., Moosbrugger, H. and Müller, H. (2003), "Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures.", *Methods of Psychological Research*, Institute for Science Education, Germany, Vol. 8 No. 2, pp. 23–74.
39. Schneider, F. (2023), "Do Different Estimation Methods Lead to Implausible Differences in the Size of Nonobserved or Shadow Economies? A Preliminary Answer", *Annual Review of Resource Economics*, Vol. 15 No. 1, pp. 257–277, doi: 10.1146/annurev-resource-090822-114308.
40. Ulyssea, G. (2018), "Firms, informality, and development: Theory and evidence from Brazil", *American Economic Review*, American Economic Association, 1 August, doi: 10.1257/aer.20141745.

41. Williams, C.C. (2008), "A critical evaluation of competing representations of the relationship between formal and informal work", *Community, Work & Family*, Vol. 11 No. 1, pp. 105–124, doi: 10.1080/13668800701785353.
42. Yusoff, M.S.B., Arifin, W.N. and Hadie, S.N.H. (2021), "ABC of Questionnaire Development and Validation for Survey Research", *Education in Medicine Journal*, Vol. 13 No. 1, pp. 97–108, doi: 10.21315/eimj2021.13.1.10.

Footnotes

1. Values as of January 5, 2026, using an exchange rate of €1 = 4,410.10 COP.

Illustration

Illustration 1 to 4 are available in the Supplementary Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterialappendix1.docx](#)
- [floatimage1.png](#)
- [floatimage2.png](#)
- [floatimage3.jpeg](#)
- [floatimage4.jpeg](#)