The Long Shade of Labor Informality

Andrés Álvarez, Oscar Becerra, Manuel Fernández*

Abstract

Countries at similar income levels exhibit markedly different rates and anatomies of labor informality. We organize these patterns around three interacting forces: a *legal wedge* (minimum wages and non-wage labor costs, alongside enforcement), the *sectoral productivity and composition*, and the *private value of formality* (coverage, portability, and contract enforceability). A parsimonious model yields sharp "thin-margin" predictions: effects concentrate where earnings cluster near the minimum legal standards. Evidence from a cross-country, country–sector panel supports the framework—legal and enforcement effects are largest where thin-margin exposure is high; higher private value lowers informality and dampens wedge effects; and composition, especially within services, conditions aggregates. The results reconcile disparate findings and imply targeted policy: align enforcement with thin-margin exposure, raise the private value of formality via low-friction administration and portability, and pursue sectoral paths that expand formal-leaning activities.

JEL Classification Codes: J46, O17, J38, J31, H55.
Keywords: Informal employment; informal sector; minimum wage; non-wage labor costs; sectoral composition.

*Álvarez: Corresponding author. Universidad de los Andes, Bogotá, Colombia, email: ca.alvarez967@uniandes.edu.co; Becerra: Universidad de los Andes, obecerra@uniandes.edu.co; Fernández: Universidad de los Andes & IZA, man-fern@uniandes.edu.co. This research was supported by the TREES Research Grant Fund. The contents of this paper are the sole responsibility of the authors and do not necessarily reflect the views of TREES. TREES is an initiative of the Center for Studies on Economic Development (CEDE) at Universidad de los Andes that promotes dialogue on inequality from the Global South through rigorous research, pedagogical innovation, and engagement with public narratives. We are also grateful to Hector Segura for his excellent research assistance.

1 Introduction

Labor informality, understood as the set of work arrangements outside the scope of government regulation, is a salient feature of labor markets in low- and middle-income economies and, in different guises, in parts of the high-income world as well (Elgin et al. 2021a). Although the share of informal work tends to decline with development, dispersion at a given income level is striking, and within-country differences across sectors are large and persistent. Any convincing account must speak both to this cross-sectional dispersion and to the internal anatomy of informality across activities, firm types, and worker contracts. This paper proposes a coherent way to organize these facts –and the research frontier around them– through a simple, empirically oriented framework.

Our starting point is that observed patterns of informality reflect the interaction of three forces shaping labor supply and demand. The first is a legal wedge on labor costs generated by statutory minimum wages and non-wage labor costs, together with the intensity and predictability of enforcement. The second is sectoral productivity and composition, which shape the selection of workers and firms into heterogeneous activities and the reallocation of employment across them. The third is the private value of formality for workers and firms—the extent to which coverage, portability, and contract enforceability raise the surplus of a formal transaction relative to an informal one. Internal heterogeneity gives these forces particular bite: economies with similar regulations can display very different aggregate outcomes because employment is distributed differently across activities where wedges bind to different degrees and where the perceived value of formality varies.

This perspective bridges classic interpretations with more recent equilibrium approaches. Dualist and structuralist traditions emphasized technology, firm size, and barriers to modern-sector expansion (e.g., Tokman 1978; Rauch 1991; Tybout 2000). Legalist views underscored non-wage costs and weak enforcement as primary drivers (Botero et al. 2004; La Porta and Shleifer 2014, among others). A newer literature models firms' and workers' choices explicitly and shows how wedges and technologies, embedded in environments with search frictions and entry costs, replicate the extensive and intensive margins observed in the data (e.g., Meghir, Narita, and Robin 2015; Ulyssea 2018). Building on these strands, we aim to deliver a small set of testable predictions that map cleanly into measurable objects, while keeping the conceptual apparatus light enough to be policy-relevant.

Two features of our approach are central. First, we place definitions and measurement before theory. Much confusion in the literature stems from mixing informal *employment*—jobs lacking social protection or legal coverage—with the *informal sector*, defined by the registration and compliance status of production units. These are distinct economic objects; conflating them blurs levels, trends, and cross-country comparisons that are central to our analysis. Second, we assemble a set of stylized facts that foreground the dispersion puzzle at a point in time, the sectoral anatomy of employment and productivity, the magnitude of wedges in their relevant context, and the key role of heterogeneity within services. Each fact is chosen to speak directly to the predictions of the model that follows and to discipline the interpretation of cross-country correlations.

The model is intentionally parsimonious. It embeds a legal floor for earnings and non-wage costs in a setting with heterogeneous activities and imperfect enforcement, and it allows the private valuation of formality to vary with the bundle of benefits and with administrative frictions that affect portability and contract enforceability. The model yields sharp comparative statics where we expect the strongest empirical bite. In particular, activities operating near the legal floor—"thin-margin" sectors—should display larger responses to enforcement or changes in the wedge, while places where the perceived value of formality is higher should exhibit lower informality at similar levels of regulation and productivity. The framework also emphasizes that aggregate changes can reflect both within-sector formalization (selection) and shifts in employment toward activities where formality is commercially viable (reallocation). These mechanisms guide an empirical program that combines panel regressions with rich fixed effects and an accounting decomposition that separates selection from reallocation components.

Our contribution is organizational rather than encyclopedic. We do not propose a new measure of the informal economy or an additional single-country case. Instead, we offer a disciplined map that reconciles influential but sometimes competing narratives, aligns measurement with international standards, and translates a fragmented literature into falsifiable claims with direct policy content. In doing so, we aim to update and extend earlier syntheses of the field (e.g., Perry et al. 2007; La Porta and Shleifer 2014; Ulyssea 2020) in light of equilibrium models of firm and worker behavior and recent evidence on structural change and services heterogeneity (e.g., Meghir, Narita, and Robin 2015; Rodrik 2016; Ulyssea 2018).

The paper proceeds as follows. Section 2 clarifies concepts and measurement of informality rates in the literature. Section 3 describes the data sources, sample construction, and key variables. Section 4 documents stylized facts that motivate the theoretical framework. Section 5 develops the model and states the principal predictions. Section 6 connects those predictions to observables and lays out the econometric strategy. Section 7 concludes.

2 Informality: Concepts, Definitions and Measures

Informality is a multidimensional concept that has been defined and measured in different ways in statistics and economics. Broadly speaking, it includes activities and practices that do not follow prescribed forms, such as those that do not meet the production and labor standards of modern economies, or established rules, such as registration with government authorities. Within that broad space, statistics office classifies the *informal sector* as those production units that are not registered or fail to comply with administrative requirements, while *informal employment* refers to jobs without legal coverage and social protection, regardless of whether they are located in formal or informal units. These distinctions, codified by the 15th and 17th International Conferences of Labor Statisticians (ICLS 1993, 2003), are not semantic: informal employment can occur inside registered firms, and some jobs in unregistered units may enjoy partial coverage (Maloney 2004; Perry et al. 2007; Ulyssea 2020; Dell'Anno 2022).

Our measurement follows this taxonomy. Where worker coverage is observed, we define informal employment as the absence of contributions/coverage in the main job (health, pensions, and related benefits), regardless of whether the employer operates in the informal sector. This definition captures non-compliance with legal mandates and is strongly correlated with other elements of the benefit package. For self-employed workers, we classify them as informal if they operate in unregistered or small household enterprises. Common proxies for informality—such as overall self-employment—are informative and we use them transparently, but they are not sufficient statistics, as they confound informality with occupational composition and survey-specific thresholds (Maloney 2004; Mondragon-Velez, Pena, and Wills 2010).

Firm-level evidence helps discipline these mechanisms. Informal firms are, on average, smaller, less capitalized, and less productive; they pay lower wages and are led by less-educated managers, even when employees' observed human capital is similar (Perry et al. 2007; La Porta and Shleifer 2008; Ulyssea 2020; Eslava et al. 2023). There is no robust "missing middle" once measurement is harmonized; rather, the prevalence of informality declines sharply with firm size and with income per capita (Perry et al. 2007; La Porta and Shleifer 2008; Paula and Scheinkman 2010). Within industries, formal firms tend to be more productive than informal peers even when aggregate productivity distributions overlap across sectors, consistent with selection on fixed costs and with barriers to scaling among informal producers (Meghir, Narita, and Robin 2015; Allen, Nataraj, and Schipper 2018; Ulyssea 2018). Over the life cycle, productivity improvements among small units are driven disproportionately by total factor productivity early on, with demand and reallocation forces becoming more important as firms grow; product-market reforms and trade exposure reallocate resources from low- to higher-productivity businesses, including movements between informal and formal margins (Eslava et al. 2004; Haltiwanger and Eslava 2017; Ulyssea 2018).

Labor-market adjustment provides a complementary lens. In downturns or under adverse shocks, informality often acts as a buffer: repressing it mechanically can raise open unemployment when modern jobs do not expand quickly enough (Dix-Carneiro et al. 2021). High worker flows in developing economies, including transitions in and out of subsistence self-employment, generate informality dynamics that differ from those in advanced economies (Donovan, Lu, and Schoellman 2023). The COVID-19 episode underscored these contrasts, with formal employment in many settings contracting more sharply at the onset and informality recovering quickly where small-scale services dominate (Alfaro, Becerra, and Eslava 2020). These patterns caution against reading informality purely as a failure of legal design; they reflect interactions between legal costs, enforcement, sectoral structure, and the private value of formal contracting.

3 Data and Sample Construction

For the empirical analysis that follows, we combine information from several harmonized international datasets. This subsection briefly describes the key sources and variable definitions; Appendix A.1 provides additional details.

Informality Rates and Employment. Data on informal employment come from ILOSTAT's SDG labor market indicators.[1] Informal employment is defined as work not covered by national labor legislation, taxation, or social protection systems, including employees without written contracts or social security coverage and self-employed workers in unregistered or small household enterprises. We use these data to construct the overall and sector-specific informality rates for each country-year. Following ILOSTAT, we group detailed industries into three broad sectors: (i) agriculture (agriculture, forestry, fishing, and related primary activities), (ii) industry (mining and quarrying, manufacturing, construction, and utilities), and (iii) services (all remaining activities, including trade, transport, accommodation and food, finance and business services, public and social services, and other services).

Using the same ILOSTAT source, we compute sectoral employment shares as the fraction of total employment in agriculture, industry, and services. Employment is defined as the number of individuals whose main activity is work contributing to the production of goods and services, irrespective of employment (employee or self-employed) or formality status. These sectoral shares are constructed to be consistent with the three-sector aggregation used for informality.

GDP per Worker and Average Earnings per Worker. Aggregate productivity is measured using two proxies from the Penn World Tables (PWT): GDP per worker and average earnings per worker. GDP per worker is defined as real GDP at constant 2017 prices (in millions of U.S. dollars) divided by total employment.

1. Indicator 8.3.1: "Proportion of informal employment in total employment by sex and sector".

Average earnings per worker is defined as total labor compensation divided by total employment. In PWT, total labor compensation equals employee compensation plus an imputed labor component for the income of the self-employed. This imputation follows the Gollin-style adjustments (Gollin 2002) implemented in PWT: when mixed-income data are available, mixed income is split between labor and capital in the same proportion as in the rest of the economy; otherwise, PWT relies on a combination of "same-wage" and agriculture-based corrections (Feenstra, Inklaar, and Timmer 2015). For employees, labor compensation includes wages and salaries in cash or in kind, plus employers' social contributions, consistent with the 2008 System of National Accounts (SNA).

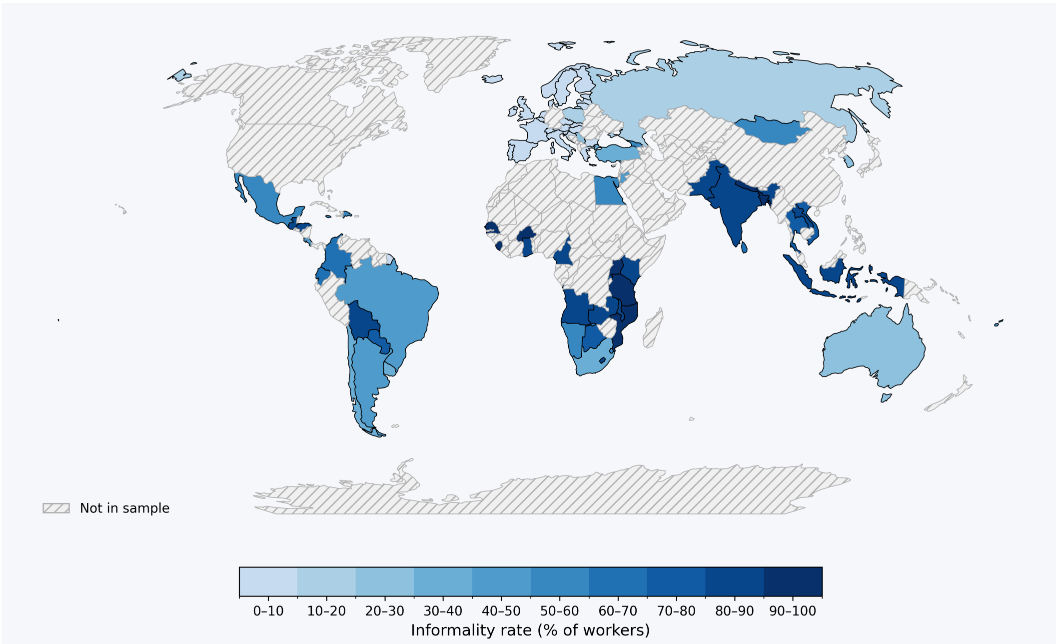
Sectoral Value Added per Worker. Sectoral labor productivity is drawn from the World Bank's Global Productivity Database (ASPD). For each country-sector-year, we compute value added per worker as real value added (in constant 2017 U.S. dollars) divided by the number of persons employed. Sectoral value added comes from national accounts statistics consistent with the 2008 SNA. The ASPD reports data for nine broad ISIC sectors; we aggregate these into agriculture, industry, and services using employment-weighted averages to match the three-sector classification used for informality and employment.

Minimum Wage. Information on statutory minimum wages is taken from ILOSTAT's series on the statutory nominal gross monthly minimum wage. This variable captures, for each country-year, the legally mandated minimum remuneration for a full month of work in the formal sector. In countries with multiple minimum wages by sector, occupation, or region, ILOSTAT reports a national benchmark rate (typically the rate applying to the largest group of workers). We convert these values to annual terms and express them in constant 2017 U.S. dollars.

Combining these sources, we construct an unbalanced panel of 81 countries for the years 2007–2019. Country and year coverage are driven primarily by the availability of labor informality data, with the aim of maximizing sample breadth while maintaining a consistent set of variables over time. Although aggregate informality rates are available for 2000–2023, coverage before 2007 is thin (at most 15 countries per year), and observations after 2019 are affected by the large and heterogeneous shocks associated with the COVID-19 pandemic. To avoid conflating these factors with medium-run trends, we focus on the 2007–2019 period. On average, countries in the sample have 7.7 years of non-missing information on informality rates. Data on sectoral value added is available only through 2017, so analyses using this variable are restricted to a 2007–2017 subsample. Data on statutory minimum wages are available for 67 of the 81 countries; the remaining 14 either lack a national statutory minimum wage or have no information on its value.

Figure 1 displays the geographical distribution of labor informality across our sample, showing country-level informality rates averaged over the full period. Informality is highest—above 70 percent of workers—in many low- and lower-middle-income economies[2] in Sub-Saharan Africa and South Asia, as well as in several Latin American countries, while most advanced economies and some upper-middle-income countries exhibit rates below 30 percent. Countries shaded with diagonal hatching are excluded from the main sample due to missing data (primarily on sectoral informality rates) and therefore do not enter the empirical analysis that follows.

Figure 1: Geographical Distribution of Labor Informality Rates



Notes: The figure plots the average informality rate (% of workers in informal employment) by country over 2007-2019. Countries are grouped into 10-percentage-point bands according to their average informality rate. Hatched areas indicate countries that are not in our main sample due to missing data on informality rates.

4 Labor Informality, Development, and Regulation

This section documents a set of empirical regularities on informal employment across countries and over time. These patterns motivate and inform the theoretical framework developed in Section 5.

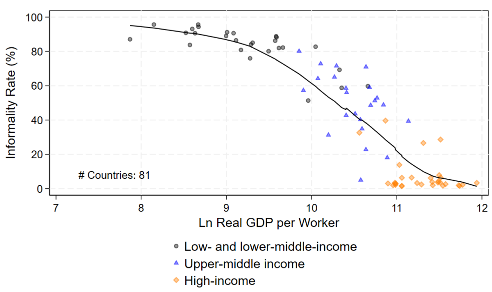
2. Countries are grouped into three income categories (low and lower-middle, upper-middle, and high) according to the World Bank's latest annual income classifications in the sample.

4.1 Development Gradient

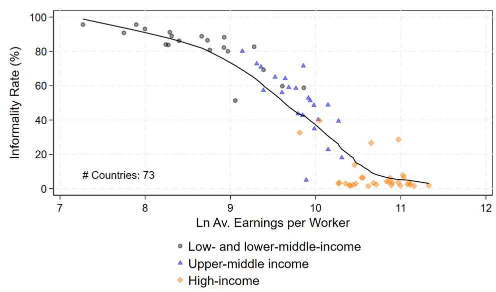
We document three regularities in the relationship between labor informality and economic development. First, informality declines sharply with development: it is highest in low- and lower-middle-income economies and falls monotonically with GDP per worker and earnings per worker (Figure 2, panels (a) and (b)). Yet countries at similar income levels display substantial dispersion in informality, indicating that the income gradient alone cannot fully explain cross-country variation. Panel (c) shows a similar pattern when income is replaced with an economy-wide human capital index, an alternative productivity proxy based on Hall and Jones (1999) and Barro and Lee (2013) (see Appendix A.1). Finally, panel (d) relates informality to self-employment, another commonly used proxy. The two indicators co-move but are not equivalent: some economies combine relatively low informality with sizable self-employment.

Figure 2: Development Gradients Across Countries

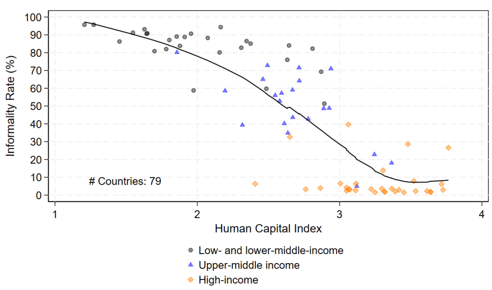
(a) Informality and GDP per Worker



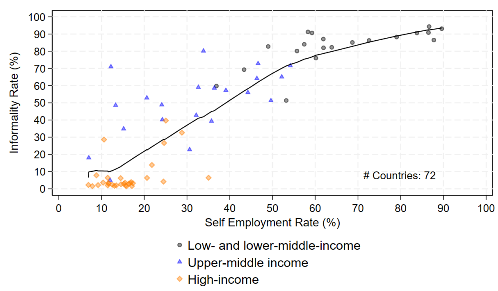
(b) Informality and Earnings per Worker



(c) Informality and Human Capital Index



(d) Informality and Self-Employment



Notes: Each marker represents the country-level average of the corresponding variables over the sample period. The sample includes all country–year observations with available data on aggregate and sectoral informality rates between 2007-2019. The number of countries with non-missing data varies by panel due to data availability. Countries are grouped into three income categories (low and lower-middle, upper-middle, and high) according to the World Bank's latest annual income classifications. The solid line in each panel depicts a locally smoothed nonparametric fit of the informality rate on the corresponding development indicator.

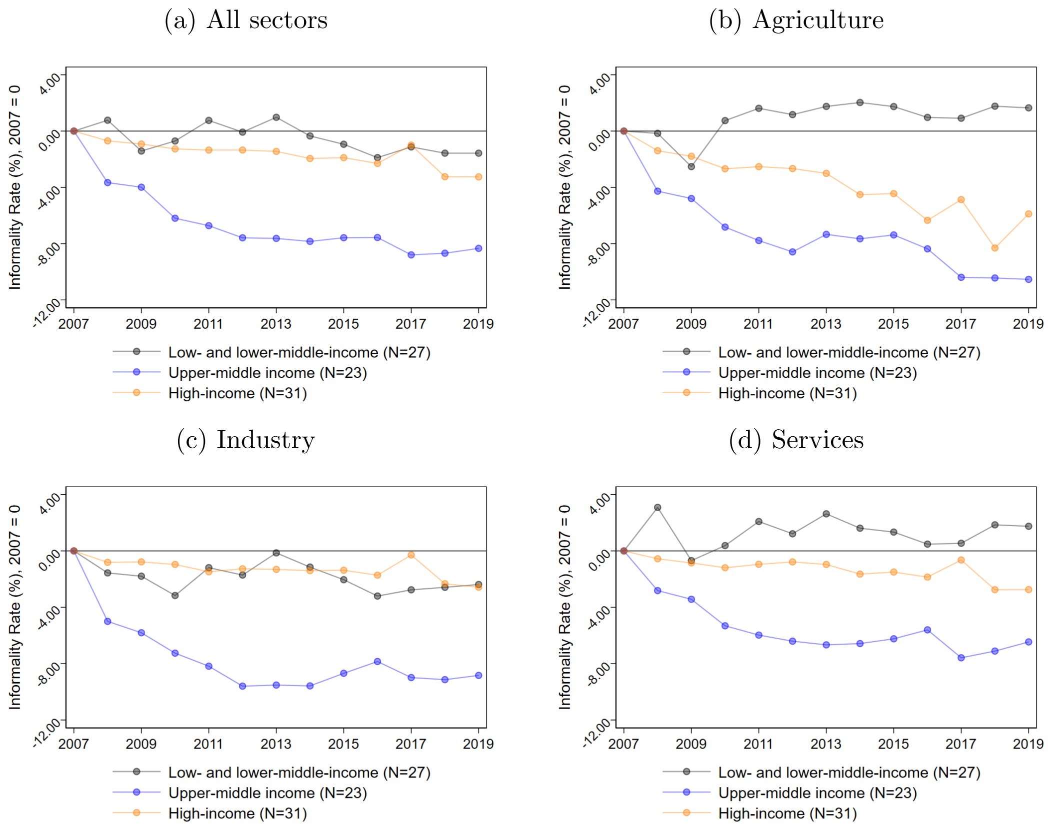
Second, within country, sectoral differences in informality rates are large and systematic. Agriculture has the highest informality rate at every income level (Figure 3, panel (a)), reaching on average 96.7% in low- and lower-middle-income economies, 75.1% in upper-middle-income economies, and a much lower 18.3% in high-income countries. Conditional on income, industry and services display substantially lower informality rates. This ranking mirrors the cross-sector productivity gradient, with value added per worker higher on average in industry and many market services than in agriculture (panel (b)). As economies develop, structural transformation reallocates employment from agriculture to industry and services (panel (c)) (Herrendorf, Rogerson, and Valentinyi 2014; Gollin and Kaboski 2023), which can mechanically reduce aggregate informality through compositional change. A further key fact also emerges: most informal workers are in services (panel (d)), but primarily because services account for a large share of total employment, not because informality rates in services are especially high.

Figure 3: Sectoral Cross-Sections by Income Group

Notes: Panels (a)–(d) report cross-country means of the corresponding variables by World Bank income group. For each income group, stacked bars decompose the aggregate value into agriculture, industry, and services; numbers above the bars report the aggregate value and the numbers in parentheses below the labels indicate the number of countries in the sample. Statistics are computed using all available country–year observations over 2007–2019, except in panel (b), where the period is 2007–2017 due to data availability for sectoral value added. Value added per worker in panel (b) is expressed in thousands of 2017 U.S. dollars. Countries are classified into income groups (low and lower-middle, upper-middle, and high) using the World Bank's latest income thresholds.

Third, over time, the time-series evidence shows that informality has declined most rapidly in upper-middle-income countries, with these declines occurring across all sectors (Figure 4). In these economies, two forces appear to be at work: falling within-sector informality and sectoral reallocation that reduces aggregate informality by shrinking agriculture's employment share (Figure 5). Reallocation away from agriculture has benefited services rather than industry, a pattern documented in prior work (Rodrik 2016). By contrast, both aggregate and within-sector declines are much smaller in low- and lower-middle-income countries, where initial informality levels were highest. In these economies, within-sector informality has remained roughly flat or increased slightly, and compositional shifts in employment across sectors have been more limited.

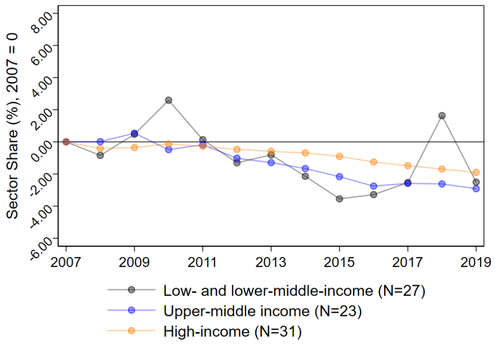
Figure 4: Sectoral Informality Trends by Income Group



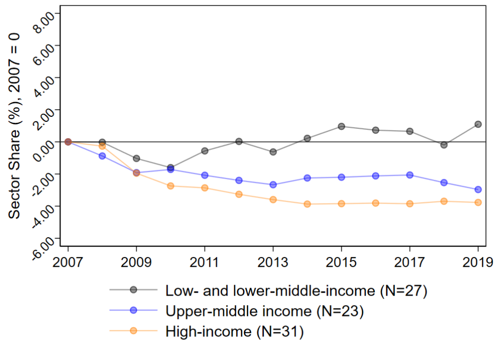
Notes: Each line plots the cross-country average informality rate for the corresponding World Bank income group over 2007-2019. The vertical axis reports deviations from the 2007 level in percentage points, so that 2007 = 0 for all series. Panel (a) shows aggregate informality, while panels (b)-(d) report informality in agriculture, industry, and services, respectively. Countries are classified into income groups (low and lower-middle, upper-middle, and high) using the World Bank's latest income thresholds.

Figure 5: Sectoral Employment Shares by Income Group

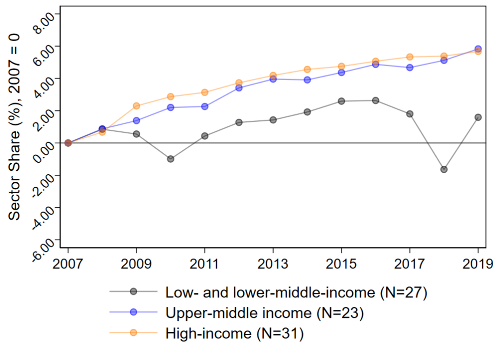
(a) Agriculture



(b) Industry



(c) Services



Notes: Each line plots the cross-country average employment share for the corresponding World Bank income group over 2007-2019. The vertical axis reports deviations from the 2007 level in percentage points, so that 2007 = 0 for all series. Countries are classified into income groups (low and lower-middle, upper-middle, and high) using the World Bank's latest income thresholds.

Taken together, these patterns line up with dual and modernization views that interpret informality as a low-productivity margin that gradually shrinks as economies grow and structural transformation proceeds (Tokman 1978; Donovan, Lu, and Schoellman 2023), while also suggest that the speed and extent of this process depend on how quickly sectors move up the productivity ladder and how workers are sorted across them.

4.2 Minimum Wage and Labor Regulations

A complementary explanation for high informality comes from the legalist and rational choice traditions. In this view, labor regulations either push formal labor costs above what a large mass of jobs can profitably sustain or, even when formal jobs remain profitable, the private value of the benefits attached to formality is too low for workers and firms to opt in when enforcement is weak. Two institutions have received particular attention in this literature: statutory minimum wages, which set a floor under formal labor costs, and non-wage labor costs arising from regulation.[3] We study both.

Figure 6 places statutory minimum wages in context. Panel (a) plots the ratio of the annualized minimum wage to GDP per worker (our wedge) against log GDP per worker, used as a proxy for aggregate productivity. A ratio of one indicates that the minimum wage equals average productivity, while lower values indicate a less binding minimum-wage floor. Panel (b) reports the same relationship using earnings per worker instead of GDP, which more closely reflects what workers are actually paid on average.

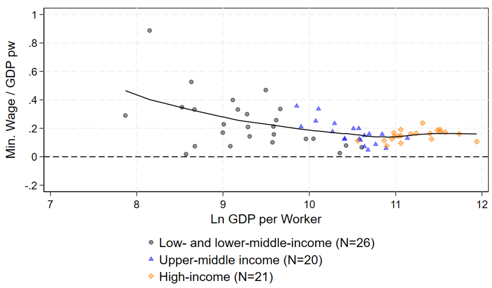
Over 2007-2019, the mean ratios are approximately 0.19 (GDP-based) and 0.37 (earnings-based), while the 90th percentiles are 0.33 and 0.70, respectively, indicating that the statutory minimum has substantial bite in many countries. In both panels, the fitted curves slope downward: in low- and lower-middle-income economies, the minimum wage tends to lie closer to average productivity, whereas in upper-middle- and especially high-income economies it is lower relative to productivity. This pattern suggests that minimum wages are more likely to bind–and thus exclude low-productivity jobs from formality–in poorer countries than in richer ones.

Panels (c) and (d) relate the aggregate informality rate to the same ratios. Consistent with this interpretation, countries with higher informality tend to display higher minimum-wage–to–productivity ratios. Among high-informality countries, however, the ratio is highly dispersed. This heterogeneity is consistent with the idea that behavior responds to the *enforced* wedge rather than the statutory one. Minimum wages matter where compliance is credible and evasion is costly; where enforcement is weak, exemptions are broad, or coverage is narrow, the same statutory floor has much less impact on informality.

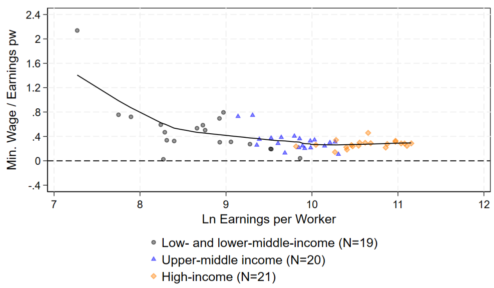
3. See, among others, Heckman et al. (2000), Almeida and Carneiro (2012), Meghir, Narita, and Robin (2015), Kugler, Kugler, and Prada (2017), Bernal et al. (2017), Ulyssea (2018, 2020), Pérez Pérez (2020), and Haanwinckel and Soares (2021).

Figure 6: Minimum Wages, Productivity, and Informality

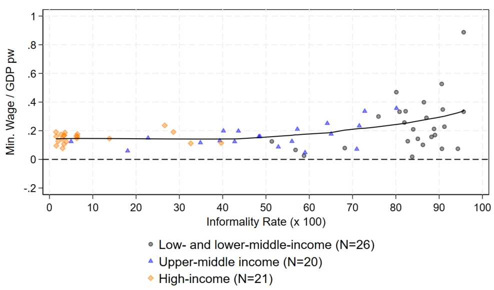
(a) Min. Wage and GDP per Worker



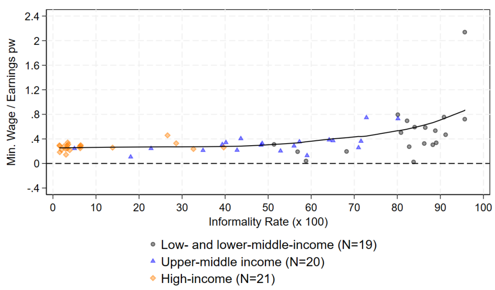
(b) Min. Wage and Earnings per Worker



(c) Min. Wage, Productivity and Informality



(d) Min. Wage, Earnings and Informality



Notes: Each marker represents a country and reports the average over 2007-2019 of the variables shown on the axes. Panels (a) and (b) plot the ratio of the (annualized) statutory minimum wage to GDP per worker and to earnings per worker, respectively, against the log of the corresponding productivity measure. Panels (c) and (d) plot the aggregate informality rate against the same ratios. The number of countries with non-missing data is 67 in panels (a) and (c), and 60 in panels (b) and (d). Countries are classified into income groups (low and lower-middle, upper-middle, and high) using the World Bank's latest income thresholds. The black curve in each panel is a locally smoothed (nonparametric) fit.

Beyond the statutory minimum wage, employers face additional non-wage costs embedded in labor regulation. To examine how the stringency of these regulations relates to informality, we draw on the *CBR Labour Regulation Index* (CBR-LRI), developed by the Center for Business Research at the University of Cambridge (Adams et al. 2023). The CBR-LRI is a comprehensive longitudinal dataset that codifies the content of national labor laws across 117 countries. It applies a leximetric methodology that converts qualitative legal provisions into quantitative indicators on a 0–1 scale, where higher values represent stronger statutory protection. Each indicator captures the formal, *de jure* scope of worker protection in specific domains, including employment protection, working-time rules, and collective labor rights. The coding relies on statutory law and case law, and, where relevant, administrative regulations and collective agreements with binding legal force. The dataset thus reflects the intended legal content of national labor legislation, rather than its enforcement or real-world implementation.

We construct an index of the *stringency of labor regulation* by aggregating 13 indicators from the CBR-LRI that capture legally mandated benefits and procedural constraints relevant to hiring and firing decisions. Specifically, the index combines:

1. *Dismissal costs for part-time workers* (CBR3);
2. *Entitlements to paid time off*, including:
    - *Annual leave* (CBR9),
    - *Public holidays* (CBR10),
    - *Overtime pay* (CBR11),
    - *Weekend pay* (CBR12);
3. *Legally mandated redundancy compensation* (CBR17);
4. *Procedural and substantive constraints on dismissal* (CBR19, CBR20);
5. *Rights to unionization and collective bargaining*, including the duty to bargain and the extension of collective agreements (CBR25–CBR28); and
6. *Rights to, and restrictions on, industrial action*, including unofficial action and the right to strike (CBR32, CBR36).

We compute the unweighted sum of these indicators for each country-year and then average it over 2007–2019 to obtain a synthetic measure. To enhance cross-country comparability and interpretability, we standardize the resulting index into z-scores, so that a one-unit corresponds to a one–standard deviation in the underlying cross-country distribution. Higher values, therefore, reflect more protective labor regulation on average.

The cross-sectional patterns are weak. Panel (a) of Figure 7 plots the labor-regulation z-score against the informality rate for the 68 countries with available data. The locally smoothed fit is essentially flat, with wide dispersion at all levels of informality: both highly regulated and lightly regulated systems appear among low- and high-informality economies. Panel (b) replaces informality with GDP per worker. The fitted relationship remains close to zero, and heterogeneity within income groups is substantial.

Several mechanisms likely account for this lack of a strong unconditional correlation. First, the CBR-LRI is a *de jure* measure: whether regulation translates into effective costs depends on enforcement capacity, inspection intensity, exemptions, and coverage, all of which vary sharply across and within countries.[4] Second, our index aggregates several distinct dimensions of labor law. These components operate through different channels: some (such as dismissal costs or payroll-related benefits)

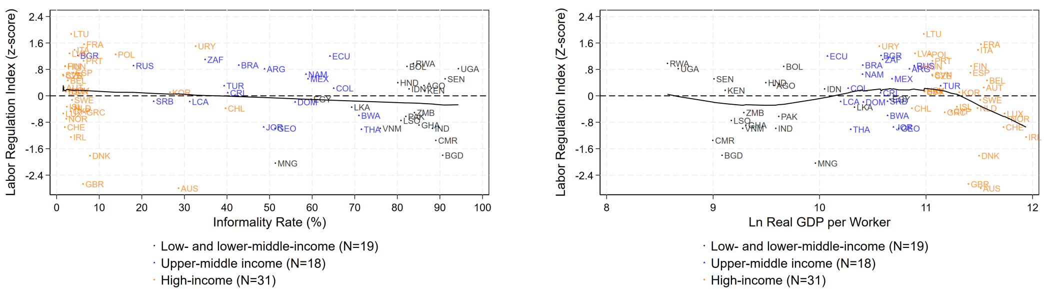
4. Micro evidence for Latin America shows that similar legal rules can have very different employment effects depending on enforcement and that weak enforcement is closely associated with the expansion of informal employment (e.g. Heckman et al. 2000; Ulyssea 2018; Haanwinckel and Soares 2021).

mainly raise marginal labor costs and may induce non-compliance or informal hiring at the margin, while others (such as collective bargaining rules) primarily affect wage setting and rents. Finally, the incidence of regulation is mediated by economic structure. Countries with similarly strict laws can differ greatly in sectoral composition, firm-size distribution, and productivity dispersion. Because enforcement tends to be stricter for larger and more visible firms, and because small, low-productivity units can often remain informal while competing with formal firms, the same *de jure* rules can generate very different mixes of formal and informal employment (Ulyssea 2018).

The near-zero cross-sectional correlations in Figure 7 should not be read as conclusive evidence that labor regulation is irrelevant for informality. Rather, they indicate that a broad *de jure* index of legal stringency captures only one layer of institutional variation and has a weak unconditional relationship with aggregate informality.

Figure 7: Informality and The Stringency of Labor Regulation

(a) Labor Regulation Index and Informality (b) Labor Regulation Index and GDP pw



Notes: Each marker reports the country-level average of the corresponding variables. The sample includes all country–year observations with available data on informality and the labor regulation index between 2007-2019. The labor regulation index is the country-average of the synthetic non-wage labor costs index, constructed from the CBR-LRI and standardized to z-scores. Panel (a) plots this index against the average informality rate; panel (b) plots it against (log) real GDP per worker over the same period. Countries are classified into income groups (low and lower-middle, upper-middle, and high-income) using the World Bank's thresholds for each year in the averaging period. The black line in each panel is a locally smoothed (nonparametric) fit. Variable definitions and data sources are described in Appendix A.1.

What do these stylized facts imply? Two elements stand out. First, shifts in productivity can reduce aggregate informality through two channels: by reallocating employment toward sectors with lower baseline informality, and by reducing within-sector informality as the productivity distribution shifts rightward. Second, there is a *wedge* channel: the minimum wage—and, potentially, non-wage labor costs—raises the profitability threshold for formal employment. When the statutory floor is close to average productivity, low-productivity matches can be excluded from the formal sector. A third *valuation* channel is also likely at work, though harder to measure directly: coverage, portability, benefits, and administrative frictions affect the private value workers and firms attach to formality. Under this channel, informality is not solely the outcome of exclusion due to low productivity under binding regulations, but also reflects an optimal choice by agents who do not value formal employment sufficiently to bear its associated costs, such as social security contributions or reduced flexibility.

5 A Parsimonious Model of Labor Informality

This section develops a parsimonious model that embeds the three forces discussed above and yields predictions that we test against the data in Section 6.

Environment. There are three sectors, indexed by $s \in \{A, M, S\}$ (agriculture, industry, services), and two work arrangements, indexed by $j \in \{F, I\}$ (formal, informal). Each worker (index omitted) is characterized by a triple of sector-specific productivities $\mathbf{y} = (y_A, y_M, y_S)$ drawn from a common joint cumulative distribution G that governs comparative advantages. This distribution can shift over time, for instance in response to neutral or sector-biased technological change.

Absent regulation, wages in sector s are a fraction $\beta \in (0, 1)$ of worker productivity, where β can be interpreted as a measure of worker's bargaining power. We assume this holds for both formal and informal work arrangements. However, employers that hire formally must pay at least the legal minimum wage w_{mw} and incur non-wage labor costs at a rate $\tau \in [0, 1)$ (e.g., payroll taxes, employer social security contributions, and other benefits). Profits from a formal match in sector s are

$$\pi_s^F(y_s) = y_s\big[1 - \beta(1+\tau)\big], \tag{5.1}$$

so non-negative profits require $\beta(1+\tau) \leq 1$. We focus on the worker-favorable case: $\beta = \frac{1}{1+\tau}$, which yields the highest bargaining weight consistent with zero-profit condition in formal jobs, i.e. the competitive benchmark.

Informal hires do not pay non-wage labor costs but face an expected enforcement penalty proportional to productivity at a rate $\varphi \in [0, 1)$, for example, because regulatory authorities concentrate enforcement on higher-value activities.[5] Expected profits from an informal match in sector s are

$$\pi_s^I(y_s) = y_s[1 - (\beta + \varphi)]. \tag{5.2}$$

Non-negative profits require $\beta + \varphi \leq 1$. Under free entry, any technology that operates in equilibrium must yield zero profits. Given $\beta = \frac{1}{1+\tau}$, this pins down the

5. This formulation captures in reduced form that both statutory penalties (often linked to the undeclared wage bill) and the probability of inspection tend to rise with the scale of the match, so that the expected cost of informality is increasing in y_s.

enforcement rate as

$$\varphi = 1 - \beta = \frac{\tau}{1+\tau}, \tag{5.3}$$

so informal matches also break even in equilibrium.

Regulation imposes a lower bound on the firm's unit labor cost under a formal hire:

$$c^F(y_s) = \max\{w_{mw}(1+\tau),\, \beta y_s(1+\tau)\} = \max\{w_{mw}(1+\tau),\, y_s\}, \tag{5.4}$$

hence, formal matches are feasible only if productivity can at least meet labor costs: $y_s \geq w_{mw}(1+\tau)$. Workers in sector s who do not satisfy this productivity threshold cannot be hired formally and are therefore *informal by exclusion.*

Finally, formality may generate non-pecuniary gains (or losses) $b \in \mathbb{R}$ for the worker (Summers 1989)—for instance, through social protection coverage and contract enforcement, net of disamenities such as reduced flexibility. We assume b scales pecuniary compensation and is drawn from a common cumulative distribution B, independent of $\mathbf{y}$. For workers who choose formal arrangements and meet productivity thresholds, total (pecuniary plus non-pecuniary) formal compensation in sector s is thus $u^F(y_s) = \beta y_s = y_s\left(\frac{1+b}{1+\tau}\right)$, whereas under informal work arrangements it is $u^I(y_s) = y_s\left(\frac{1}{1+\tau}\right)$.

Worker's Choice. Because τ, w_{mw}, and a worker's b are the same across sectors, the worker's total compensation in sector s is a strictly increasing function of y_s, so the sector that maximizes compensation is also the one with the highest productivity. We therefore model sectoral choice in a two-step Roy-style framework. First, the worker chooses the sector in which her productivity is highest; second, conditional on that sector, she decides whether to work formally or informally if her productivity is sufficiently high and her private valuation of formality is positive.

Sectoral employment shares are given by

$$\theta_s = \Pr\Big(s = \arg\max_{k\in\{A,M,S\}} y_k\Big), \qquad \text{with} \sum_{s\in\{A,M,S\}} \theta_s = 1. \tag{5.5}$$

Conditional on working in sector s, a worker chooses formal employment if

$$u^F(y_s) \geq u^I(y_s) \iff b \geq 0 \quad \text{and} \quad y_s \geq w_{mw}(1+\tau).$$

The informality rate by *exclusion* in sector s is

$$\mathrm{Inf}_x(s) = \Pr\left(y_s < w_{mw}(1+\tau) \mid s = \arg\max_{k\in\{A,M,S\}} y_k\right). \tag{5.6}$$

Among workers in sector s who are not excluded (i.e., $y_s \geq w_{mw}(1+\tau)$), a fraction

$B(0)$ chooses informality. The informality rate by *choice* in sector s is therefore

$$\text{Inf}_c(s) = \Big(1 - \text{Inf}_x(s)\Big)B(0), \tag{5.7}$$

while the sector-specific informality rate is

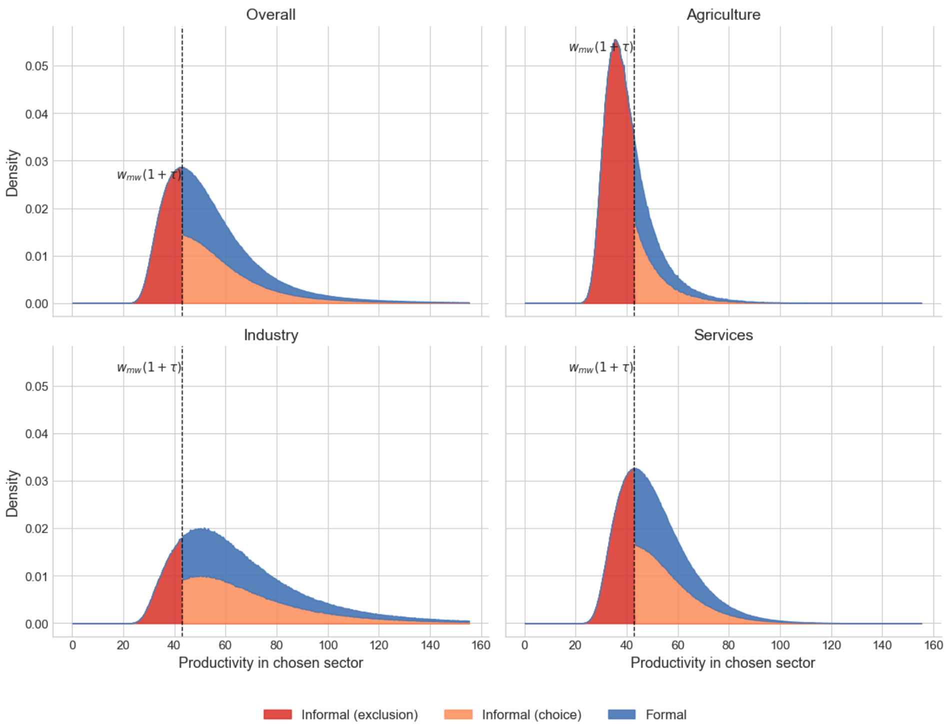
$$\begin{aligned} \text{Inf}(s) &= \text{Inf}_x(s) + \text{Inf}_c(s) \\ &= B(0) + \Big(1 - B(0)\Big)\text{Inf}_x(s). \end{aligned} \tag{5.8}$$

Finally, aggregate informality is the sector-share-weighted average

$$\text{Inf} = \sum_s \theta_s \, \text{Inf}(s). \tag{5.9}$$

Figure 8 displays simulated productivity distributions by sector and work arrangement. In the simulations, sector-specific productivities y_s follow log-normal distributions, while non-pecuniary valuations of formality b are drawn from a standard uniform distribution. The stacked densities decompose total employment into three groups: informal by exclusion (workers with productivity below the minimum-wage threshold $w_{mw}(1+\tau)$), informal by choice, and formal. Informality by exclusion is concentrated in the lower tail; as productivity rises beyond the threshold, informality declines and becomes increasingly driven by choice. In line with the literature, we obtain informal employment among both low- and high-productivity matches within the same sector (Ulyssea 2020).

Figure 8: Simulated Productivity Distributions by Sector and Work Arrangement



Notes: Each panel plots the simulated density of productivity in the worker's chosen sector, decomposed into three groups: informal by exclusion (red), informal by choice (orange), and formal (blue). Sectoral productivities y_s are log-normally distributed. The non-pecuniary valuation of formality b is drawn from a uniform distribution on $[-1, 1]$. Workers are informal by exclusion if their productivity in the chosen sector falls below the minimum-wage productivity threshold $w_{mw}(1+\tau)$; among eligible workers, those with $b < 0$ are informal by choice and those with $b \geq 0$ are formal. The top-left panel shows the aggregate distribution across all sectors; the remaining panels show the corresponding distributions conditional on each sector.

Comparative Statics. Four sets of predictions follow directly.

1. Dynamics. Small changes in informality decompose into reallocation across sectors and selection within sectors (intensive margin):

$$d\mathrm{Inf} = \underbrace{\sum_s d\theta_s \,\mathrm{Inf}(s)}_{\text{reallocation across sectors}} + \underbrace{\sum_s \theta_s \, d\mathrm{Inf}(s)}_{\text{selection within sectors}} . \tag{5.10}$$

This mirrors the accounting in Section 4: sectoral productivity shifts move θ_s and $\mathrm{Inf}(s)$; changes in w_{mw} and τ move $\mathrm{Inf}(s)$ through exclusion; shifts in B move $\mathrm{Inf}(s)$ through the choice margin.

2. Sectoral Productivity and Composition. A favorable sectoral (or aggregate) productivity shock that shifts up the distribution of y_s in a formal-leaning activity reduces informality through two channels. First, *selection*: more matches in that sector cross the productivity cutoff, lowering $\mathrm{Inf}_x(s)$ and hence $\mathrm{Inf}(s)$. Second, *reallocation*: higher expected payoffs raise θ_s, shifting employment toward a lower-informality sector and reducing aggregate informality (Inf). In this sense, sector-biased (non-neutral) technical change that disproportionately raises productivity in formal-leaning activities amplifies the decline in informality.

Structural change out of agriculture—the sector with the highest informality—is a salient case: as productivity in non-agricultural activities improves, employment reallocation away from agriculture and toward industry and services lowers aggregate labor informality both by shrinking a highly informal sector and by increasing the weight of less informal ones. The aggregate impact scales with sectoral employment weights, consistent with the cross-country patterns in Section 4. Medium-run outcomes depend on how quickly productivity and employment reallocate across sectors.

3. The Minimum Wage Wedge. The minimum wage and non-wage labor costs affect informality only through the exclusion threshold. Let $T \equiv w_{mw}(1+\tau)$ denote the minimum-wage productivity cutoff. Differentiating (5.8) with respect to T,

$$\frac{\partial \text{Inf}(s;T)}{\partial T} = \big(1 - B(0)\big)\frac{\partial \text{Inf}_x(s;T)}{\partial T} = \big(1 - B(0)\big)\, h_s(T) > 0, \tag{5.11}$$

where $h_s(T)$ is the density of y_s at the cutoff. Thus a higher minimum wage or higher non-wage labor costs (a larger T) increase informality by expanding the mass of workers who are informal by exclusion, partially offset by a decline in informality by choice. Importantly, the impact is larger in sectors where many matches are close to the threshold—that is, where $h_s(T)$ is high, typically low-value-added activities. Since T does not affect productivities $\mathbf{y}$, sectoral employment shares θ_s are unchanged: the wedge channel operates entirely through the formal–informal margin, not through sectoral reallocation.

At the aggregate level,

$$\text{Inf}(T) = \sum_s \theta_s \text{Inf}(s;T) \quad \Rightarrow \quad \frac{\partial \text{Inf}(T)}{\partial T} = \sum_s \theta_s \big(1 - B(0)\big)\, h_s(T) > 0. \tag{5.12}$$

Hence, increases in w_{mw} or τ unambiguously raise aggregate informality by shifting more low-productivity jobs into exclusion, while leaving sectoral employment shares θ_s unchanged.

4. Private Value of Formality. Holding technology and policy fixed, an upward shift in the distribution B that raises the valuation of formality reduces informality through the choice margin. Using (5.8)–(5.9), the change in aggregate informality between two distributions B_0 and B_1 is

$$\Delta \text{Inf} = \sum_s \theta_s \Big(\text{Inf}(s;B_1) - \text{Inf}(s;B_0)\Big) = \sum_s \theta_s \big(1 - \text{Inf}_x(s)\big)\big(B_1(0) - B_0(0)\big) < 0. \tag{5.13}$$

Because a shift that makes formality more attractive lowers the mass of workers with $b < 0$, i.e. $B_1(0) < B_0(0)$. The effect is largest in sectors where many workers are eligible for formal employment ($1 - \text{Inf}_x(s)$ is high) and where a sizable share of workers is close to indifference between formal and informal arrangements.

This mechanism also shapes how productivity shocks translate into changes in informality. From (5.6)–(5.8), the impact of a productivity shift that relaxes exclusion in sector s is proportional to $1 - B(0)$, the share of workers who weakly prefer formality. Hence, for a given improvement in productivity, informality falls more in economies where the baseline valuation of formality is higher (lower $B(0)$), and less in economies where a large mass of workers prefers informal arrangements.

6 From Theory to Evidence: Econometric Exercises

In this section, we test the model's predictions against the data. The evidence is correlational and complements that presented in Section 4, but here we attempt to isolate the predicted mechanisms within our cross-country setting.

6.1 Dynamics

We begin with the decomposition of aggregate informality into reallocation and selection in Equation 5.10. The discrete-time counterpart for country c between years t and t' is

$$\Delta\text{Inf}_c = \underbrace{\sum_s \Delta\theta_{c,s}\,\text{Inf}_{c,s,t}}_{\text{reallocation across sectors}} + \underbrace{\sum_s \theta_{c,s,t}\,\Delta\text{Inf}_{c,s}}_{\text{selection within sectors}} + \underbrace{\sum_s \Delta\theta_{c,s}\,\Delta\text{Inf}_{c,s}}_{\text{interaction term}}, \tag{6.1}$$

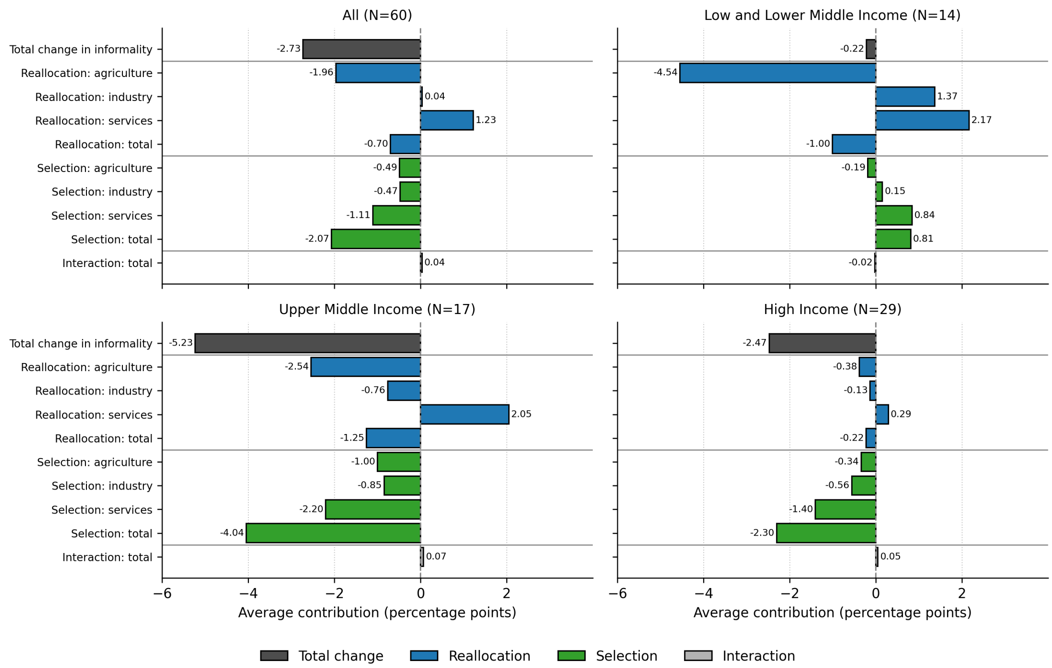
where $\theta_{c,s,t}$ and $\text{Inf}_{c,s,t}$ denote initial sectoral employment shares and informality rates, and Δ denotes the change between t and t'. The first term captures changes in aggregate informality coming purely from movements in sectoral composition, holding within-sector informality fixed at its initial level. The second term captures changes coming purely from shifts in within-sector informality, holding sectoral employment shares fixed at their initial level. The third term is an interaction term that reflects the fact that both sectoral shares and within-sector informality can change simultaneously.

Since our panel is unbalanced, the initial and final years (t_c, t'_c) vary across countries. We implement the decomposition for the 60 countries that have at least two observations on informality rates separated by four or more years: $\Delta_c = t'_c - t_c \geq 4$. In this sample, the average gap is 10.1 years and the 10th percentile is 7 years. To make the results comparable across countries, we rescale $\Delta\theta_{c,s}$ and $\Delta\text{Inf}_{c,s}$ by multiplying each change by the factor $10/\Delta_c$, so that all components of ΔInf_c are interpreted on a per-decade basis.

Figure 9 reports the decomposition results. Two main patterns emerge. First, pulling all income groups, the selection component is quantitatively more important than reallocation: it accounts for roughly three quarters of the decline in informality, while reallocation accounts for most of the remaining fall and the interaction term is negligible. Within the selection margin, the largest reductions occur in services, whereas on the reallocation margin the largest contribution comes from shifts out of agriculture.

Second, there is marked heterogeneity by income group. Among low- and lower-middle-income countries, reallocation away from agriculture exerts a sizeable downward force on aggregate informality, but this is almost fully offset by rising within-sector informality in industry and services, so the net change is close to zero. The combination of these two forces suggests that many of the new jobs created outside agriculture remain informal: workers may experience earnings gains by switching sectors, but are still sorted into informal work arrangements. By contrast, in upper-middle- and high-income countries, within-sector informality declines in all sectors, with reallocation playing a smaller but non-trivial role. This pattern is more consistent with broad-based productivity gains in non-agricultural sectors that reduce informality even when movers are not positively selected on productivity.

Figure 9: Decomposition of Changes in Informality by Income Group



Notes: The figure reports the decomposition of the change in aggregate informality between the first and last year observed for a sample of 60 countries with at least two observations on informality separated by four or more years. Changes in the informality rate are decomposed into reallocation (between-sector), selection (within-sector), and interaction components across agriculture, industry, and services (see Equation (6.1)). Bars show the average contribution of each component (in percentage points per decade) across countries within each World Bank income group.

.

6.2 Productivity and Composition

We now examine how our aggregate productivity proxies relate to informality rates and sectoral employment shares. The model predicts that, holding other determinants constant, improvements in aggregate productivity should unambiguously reduce informality and, when sector-biased, trigger a reallocation of workers toward the sectors that experience the largest productivity gains.

As a starting point, we estimate the model

$$Y_{c,t} = \mu_c + \delta_{r(c),t} + \psi_{i(c),t} + \tau \text{Log GDP\_pw}_{c,t} + e_{c,t}, \tag{6.2}$$

where $Y_{c,t}$ denotes the outcome of interest (informality or employment shares) in country c and year t; μ_c are country fixed effects; $\delta_{r(c),t}$ are region-by-year fixed effects,[6] $\psi_{i(c),t}$ are income-group-by-year fixed effects,[7] and $e_{c,t}$ is the error term.

The coefficient τ captures the association between log GDP per worker and the outcome, net of time-invariant country heterogeneity and common shocks specific to region and income group. Thus, identification comes from productivity changes that deviate from the average evolution of countries in the same region and income group. As a robustness check, we re-estimate (6.2) replacing Log $\text{GDP\_pw}_{c,t}$ with Log $\text{Earnings\_pw}_{c,t}$, which may better capture productivity gains accruing directly to labor rather than to capital.

Results are presented in Table 1. Consistent with the development gradient documented in Section 4, countries that experience above-average increases in GDP per worker display lower informality rates (Column 1). The point estimate implies that a 1% increase in GDP per worker is associated with a 0.169 percentage-point decline in the aggregate informality rate. Using earnings per worker instead of GDP yields very similar semi-elasticities, suggesting that the patterns are not driven by changes in the capital share. These estimates likely represent a lower bound on the effect of productivity on informality, since in many settings minimum-wage adjustments are at least partly linked to productivity growth, dampening the estimated reduction—a point to which we return in the next section.

Columns (2)–(4) show that these relationships are concentrated in non-agricultural sectors. The semi-elasticities for industry and services are sizeable and statistically significant, whereas the coefficient for agriculture is close to zero and imprecisely estimated. This is not because agriculture experiences weaker productivity growth in proportional terms: Figure 10 shows that value added per worker increases the most in agriculture over our sample period. Rather, average productivity in agriculture

6. Regions follow the World Bank classification and include East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, South Asia, and Sub-Saharan Africa.

7. Income groups follow the World Bank classification and include low- and lower-middle-income, upper-middle-income, and high-income countries.

is much lower to begin with (Panel (b) of Figure 3), so a given percentage increase implies a relatively small absolute gain. If most agricultural workers lie well below the exclusion threshold for formality, even substantial proportional improvements in productivity will not move many of them across that margin.

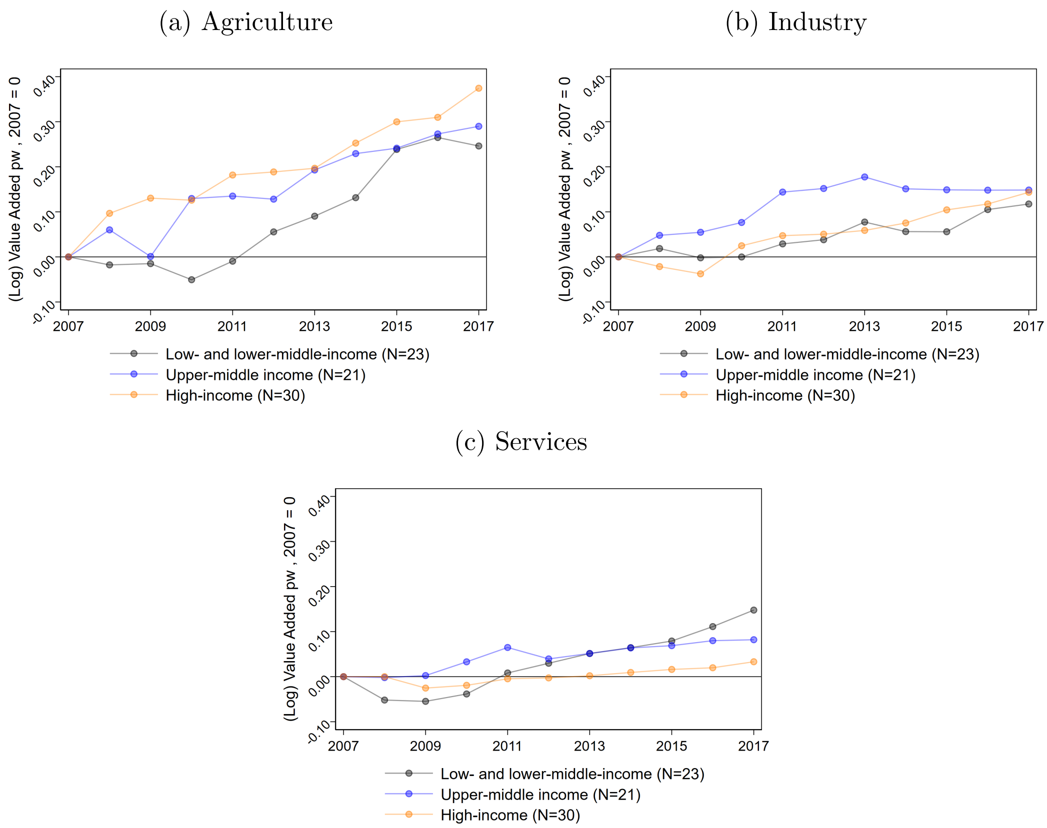
Columns (5)–(7) examine sectoral employment shares. In line with the structural change literature, higher GDP per worker is associated with a reallocation of employment out of agriculture: the agricultural employment share falls, while the industrial share rises, with no statistically significant change in services. The results using earnings per worker display the same qualitative pattern, albeit with somewhat smaller magnitudes and an imprecisely estimated effect on agriculture. Overall, these correlations are consistent with a joint process of structural transformation and declining within-sector informality in industry and services as aggregate productivity rises.

Table 1: Informality and Employment Shares by Aggregate Productivity

	Informality Rate				**Employment Share**		
	All Sectors (1)	Agriculture (2)	Industry (3)	Services (4)	Agriculture (5)	Industry (6)	Services (7)
Log GDP per Worker	-0.169**	-0.013	-0.169***	-0.142**	-0.089*	0.067**	0.024
	(0.056)	(0.110)	(0.041)	(0.059)	(0.048)	(0.025)	(0.035)
Observations	622	622	622	622	622	622	622
Countries	81	81	81	81	81	81	81
Mean of Dep. Var. (%)	33.1	50.9	30.5	27.9	15.5	22.5	62.0
Log Earnings per Worker	-0.142**	-0.021	-0.125***	-0.173***	-0.005	0.045**	-0.040
	(0.049)	(0.097)	(0.033)	(0.047)	(0.030)	(0.019)	(0.025)
Observations	588	588	588	588	588	588	588
Countries	73	73	73	73	73	73	73
Mean of Dep. Var. (%)	30.5	48.3	27.9	25.8	13.7	22.7	63.5

Notes: Each column reports coefficients from regressions of the indicated outcome on measures of labor productivity. Panel A uses log GDP per worker; Panel B uses log earnings per worker. All specifications include country fixed effects and year fixed effects interacted with region and World Bank income group. Standard errors are clustered at the country level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 10: Log Value Added per Worker by Income Group and Sector



Notes: Each line plots the cross-country average (log) value added per worker for the corresponding World Bank income group over 2007–2017. The vertical axis reports deviations from 2007, so that 2007 = 0 for all series.

Shift-Share. To address the potential endogeneity of our productivity measures, and to more closely mirror the sectoral-share structure of informality in the model, we complement the panel regressions with a shift–share analysis on the subsample of countries observed in both 2007 and 2017, the first and last years for which sectoral value added is available.

We first collapse the panel to a country-level long difference between 2007 and 2017. The change in aggregate informality is

$$\Delta \text{Inf}_c \equiv \text{Inf}_{c,2017} - \text{Inf}_{c,2007}. \tag{6.3}$$

Let $\theta_{c,s,2007}$ denote the initial employment share of sector s in country c, and let

$$\Delta \text{Log VA\_pw}_{c,s} \equiv \text{Log VA\_pw}_{c,s,2017} - \text{Log VA\_pw}_{c,s,2007} \tag{6.4}$$

denote the corresponding change in (log) value added per worker in sector s. We then construct a country-specific ("internal") shift–share index,

$$Z_c = \sum_s \theta_{c,s,2007}\,\Delta\text{Log VA\_pw}_{c,s}, \tag{6.5}$$

which summarizes within-country sectoral productivity growth, weighted by the country's initial employment structure.

To separate idiosyncratic country shocks from common shocks at a given level of development, we also build an "external" shift–share index using the same base-year employment weights $\theta_{c,s,2007}$, but replacing $\Delta\text{Log VA\_pw}_{c,s}$ with the average sectoral productivity change among other countries in the same income group $i(c)$:

$$Z_c^{\text{ext}} = \sum_s \theta_{c,s,2007}\,\Delta\text{Log VA\_pw}_{i(c),s}, \tag{6.6}$$

where

$$\Delta\text{Log VA\_pw}_{i(c),s} = \frac{1}{N_{i(c)}-1}\sum_{\{j:j\neq c,\, j\in i(c)\}} \Delta\text{Log VA\_pw}_{j,s},$$

and $N_{i(c)}$ is the number of countries in income group $i(c)$. By construction, Z_c^{ext} is driven by productivity shocks that are common to countries at a similar income level, while excluding country c's own sectoral shocks, making it less likely that the index is contaminated by country-specific confounders.

We then estimate cross-country regressions of the form

$$\Delta\text{Inf}_c = \delta_{r(c)} + \psi_{i(c)} + \beta Z_c + \epsilon_c, \tag{6.7}$$

and analogously replacing Z_c with Z_c^{ext}, where $\delta_{r(c)}$ and $\psi_{i(c)}$ are fixed effects for World Bank region and income group, and ϵ_c is an error term.

For ease of interpretation, both shift–share indices are standardized to have mean zero and unit standard deviation across countries. The coefficient β therefore measures the change in the aggregate informality rate associated with a one–standard-deviation increase in the corresponding shift–share index.

Results are reported in Table 2. The estimates are qualitatively similar to the panel regressions, despite the smaller country sample. In our preferred specification (Column 4), a one–standard-deviation increase in the external shift–share index is associated with a 3.1 percentage-point decline in aggregate informality over the decade, which corresponds to roughly 75% of the observed average change in the sample. These results reinforce the view that productivity dynamics are a central determinant of labor informality in the medium term.

Table 2: Changes in Informality and Shift–Share Indices of Sectoral Productivity Growth

	Informality Rate			
	(1)	(2)	(3)	(4)
Shift Share (Z_c)	-0.021*	-0.022		
(Standardized)	(0.011)	(0.013)		
Shift Share (Z_c^{ext})			-0.025***	-0.031**
(Standardized)			(0.006)	(0.015)
Observations	26	26	26	26
Countries	26	26	26	26
Mean of Dep. Var. (%)	-4.1	-4.1	-4.1	-4.1
Region FE	No	Yes	No	Yes
Income-Group FE	No	Yes	No	Yes

Notes: Each column reports cross-country regressions of the change in the aggregate informality rate between the first and last year observed (2007–2017) on shift–share indices of sectoral value added per worker growth. The dependent variable is the difference in the informality rate between the initial and final year in percentage points. "Shift Share (Z_{ct})" is constructed using country-specific base-year employment shares by sector (agriculture, industry, services) and within-country changes in log value added per worker by sector. "Shift Share (Z_{ct}^{ext})" is constructed using the same base-year employment shares but replacing within-country sectoral productivity changes with averages for other countries in the same World Bank income group. Both shift–share indices are standardized to have mean zero and unit standard deviation. Standard errors estimated using robust standard errors.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.3 The Minimum Wage Wedge

We now turn to the model's predictions regarding minimum-wage changes: (i) increases in the statutory minimum wage should raise aggregated informality, and (ii) they should have no first-order effect on sectoral employment shares. The first prediction is state-dependent: minimum-wage hikes are predicted to have larger effects when the minimum wage has "bite"—that is, when it is close to the earnings of a large mass of workers. Moreover, because minimum wages are often adjusted in response to productivity changes, what matters empirically is not the minimum-wage level itself, but how the minimum-wage to earnings wedge evolves over time and whether the impact of these changes differs by the initial wedge.

To test the first prediction on aggregate informality, we estimate

$$\begin{aligned}\text{Inf}_{c,t} =&\mu_c + \delta_{r(c),t} + \psi_{i(c),t} + \tau_1 \text{Log}\left(\frac{\text{MW}_{c,t}}{\text{Earnings\_pw}_{c,t}}\right) \\ &+ \tau_2 \text{Low\_Wedge}_c \times \text{Log}\left(\frac{\text{MW}_{c,t}}{\text{Earnings\_pw}_{c,t}}\right) + e_{c,t},\end{aligned} \tag{6.8}$$

where $\text{Log}\left(\frac{\text{MW}_{c,t}}{\text{Earnings\_pw}_{c,t}}\right)$ is the (log) wedge, so higher values indicate a tighter (more binding) minimum wage relative to labor productivity; μ_c are country fixed effects; $\delta_{r(c),t}$ are region-by-year fixed effects; and $\psi_{i(c),t}$ are income-group-by-year fixed effects. In some specification, we include log GDP per worker as a control.

The dummy Low_Wedge_c is an indicator for high-bite countries, equal to one for those whose time-averaged minimum-wage–to–earnings ratio lies above the 90th percentile of the sample distribution, which in our data corresponds to a ratio of approximately 0.7. In this specification, τ_1 captures the semi-elasticity of informality with respect to the wedge in the baseline group, while $\tau_1 + \tau_2$ gives the corresponding semi-elasticity for high-bite countries.

Results are presented in Table 3. Two main patterns emerge. First, for countries in the baseline group ($\text{Low\_Wedge}_c = 0$), changes in the wedge have no statistically detectable effect on aggregate or sectoral informality: the coefficients on (log) Wedge are small and imprecise in all columns. Second, in high-bite countries ($\text{Low\_Wedge}_c = 1$), the interaction term is positive and, for aggregate informality as well as for industry and services, statistically significant. The implied semi-elasticity $\tau_1 + \tau_2$ for aggregate informality lies between 4.3 and 7.8 percentage points, which corresponds to an effect equal to about 12–23% of the sample mean informality rate.

Sectorally, the strongest effects occur in services, where the corresponding semi-elasticities range from 6.2 to 9.8 percentage points (roughly 22–34% of the sample mean), followed by industry (5.2–8.5 percentage points, or about 16–27% of the sample mean), while agriculture shows no systematic response. The absence of effects in agriculture is consistent with the theoretical framework: because average productivity in agriculture is comparatively low, many agricultural workers are below the threshold regardless of the exact level of the minimum wage, so the wedge is largely irrelevant for their formality decision.

Overall, the evidence suggests that minimum-wage hikes can have adverse effects on informality, but not unconditionally: their impact depends critically on the level of the minimum wage relative to average earnings.

Table 3: Informality and the Minimum-Wage to Earnings Wedge

	Informality Rate: All Sectors				**Informality Rate by Sector**					
					Agriculture	Agriculture	Industry	Industry	Services	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Log Wedge	0.016	0.018*	0.011	0.016	-0.005	-0.005	-0.002	0.002	0.011	0.016
	(0.011)	(0.011)	(0.010)	(0.011)	(0.021)	(0.021)	(0.007)	(0.007)	(0.012)	(0.013)
(2) Log Wedge X Low Wedge			0.067**	0.027	0.008	0.010	0.087***	0.050*	0.088**	0.046
			(0.020)	(0.023)	(0.049)	(0.051)	(0.021)	(0.025)	(0.026)	(0.029)
Log GDP per Worker		-0.178**		-0.169**		0.009		-0.156***		-0.174**
		(0.059)		(0.061)		(0.125)		(0.043)		(0.060)
Sum Coefs. (1)+(2)			0.078***	0.043**	0.003	0.005	0.085***	0.052**	0.098***	0.062**
P-Val.			0.000	0.042	0.949	0.921	0.000	0.030	0.000	0.023
Observations	484	484	484	484	484	484	484	484	484	484
Countries	60	60	60	60	60	60	60	60	60	60
Statistics:										
Mean of Dep. Var. (%)	33.9	33.9	33.9	33.9	53.3	53.3	31.0	31.0	28.8	28.8
Mean MW–to–Earnings (%) (*High-Wedge Countries*)	28.1	28.1	28.1	28.1	28.1	28.1	28.1	28.1	28.1	28.1
Mean MW–to–Earnings (%) (*Low-Wedge Countries*)	79.5	79.5	79.5	79.5	79.5	79.5	79.5	79.5	79.5	79.5

Notes: Each column reports coefficients from regressions of informality rates on the ln minimum-wage to earnings wedge. Columns (1)–(4) use overall informality (share of informal workers in total employment) as the dependent variable. Columns (5)–(10) use sectoral informality (share of informal workers in sectoral employment) for agriculture, industry, and services. *Log Wedge* is the log ratio of the statutory minimum wage to average earnings per worker. *Log Wedge* $\times$ *Low Wedge* interacts this term with an indicator equal to one for countries whose average minimum-wage–to–earnings ratio lies above the 90th percentile of the sample distribution (approximately 0.70). All specifications include country fixed effects and year fixed effects interacted with region and World Bank income group. Standard errors are clustered at the country level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The second prediction is that minimum-wage hikes, by themselves, should not alter the allocation of employment across sectors. Table 4 provides evidence consistent with this prediction. We estimate a model analogous to equation (6.2), using sectoral employment shares as dependent variables and $\text{Log}\,\text{MW}_{c,t}$ as the main regressor, while including the same set of country, region-by-year, and income-group–by–year fixed effects. In some specifications we also control for $\text{Log}\,\text{Earnings\_pw}_{c,t}$.

Across all sectors and specifications, the estimated effects of the minimum wage on employment shares are small and statistically indistinguishable from zero, indicating no systematic reallocation of employment across agriculture, industry, and services in response to minimum-wage changes.

Table 4: Minimum Wages and Sectoral Employment Shares

	Employment Share by Sector					
	Agriculture (1)	Agriculture (2)	Industry (3)	Industry (4)	Services (5)	Services (6)
Log Min. Wage	0.016	0.019	-0.000	-0.007	-0.016	-0.011
	(0.017)	(0.019)	(0.008)	(0.007)	(0.013)	(0.015)
Log Earnings per Worker		-0.021		0.056**		-0.036
		(0.041)		(0.021)		(0.033)
Observations	484	484	484	484	484	484
Countries	60	60	60	60	60	60
Mean of Dep. Var. (%)	14.9	14.9	22.5	22.5	62.5	62.5

Notes: Each column reports coefficients from regressions of sectoral employment shares on the statutory minimum wage. The dependent variable is the employment share of the indicated sector (agriculture, industry, services) in total employment, expressed in percentage points. *Log Min. Wage* is the log of the statutory minimum wage in real terms. The sample is restricted to country–year observations for which both sectoral employment and the minimum-wage–to-earnings ratio are available. All specifications include country fixed effects and year fixed effects interacted with region and World Bank income group. Standard errors are clustered at the country level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.4 Private Value of Formality

The last prediction of the model concerns the private value of formality and its role in shaping individual choices over formal work arrangements. Empirically, informality is not confined to the bottom of the productivity or earnings distribution; it is observed throughout (Maloney 2004; Meghir, Narita, and Robin 2015; Ulyssea 2018, 2020). From a rational-choice perspective, this is consistent with *informality by choice*: conditional on being productive enough to qualify for a formal job, some workers and firms optimally choose informality when the perceived costs of formality (taxes, contributions, regulation) exceed the private value of the associated benefits (insurance, protection, job safety) (Levy 2008). In the model, the decision to work formally, conditional on clearing the exclusion threshold, is governed by a cost–benefit comparison summarized by the distribution of the private valuation term b.

We do not observe either the distribution of preferences for formality or the intensity of enforcement and penalties for non-compliance, so we take an indirect approach. Equations (5.8) and (5.9) imply that positive productivity shocks reduce informality by shrinking the mass of workers who are excluded by the minimum-wage–productivity threshold. However, the magnitude of this effect depends on how many of the newly eligible workers actually choose formal jobs, which in turn depends on the private value of formality. In other words, the elasticity of informality with respect to productivity should be larger in environments where workers place a higher value on the attributes of formal employment.

We approximate these preferences using information from the World Values Survey (WVS). Specifically, we exploit the question: *"Regardless of whether you're actually looking for a job, which one would you, personally, place first if you were looking for a job?"* One of the response options is *"A safe job with no risk of closure or unemployment."* We interpret the share of respondents choosing this option as their top priority as an imperfect but relevant proxy for the private valuation of formality: formal jobs are more likely to offer legal protection, dismissal regulation, and access to social insurance, whereas informal jobs, by virtue of being unregulated, tend to be more precarious and exhibit larger cyclical fluctuations (Eslava et al. 2004; Haltiwanger and Eslava 2017; Ulyssea 2018).

This question is available for a subsample of 23 countries in our main sample, but only in WVS waves between 1981 and 2003. For each country–wave, we compute the percentage of respondents who report *"A safe job with no risk of closure or unemployment"* as their top choice and then average across waves to obtain a time-invariant country-level measure. We treat this average as a long-run proxy for the country's valuation of job security at baseline. In our sample, an average of 36.5% of respondents choose job security as their first priority, with a standard deviation of 8.7 percentage points across countries.

We standardize the measure to have mean zero and unit standard deviation across countries and use it as a moderator in the regression

$$\begin{aligned} \text{Inf}_{c,t} &= \mu_c + \delta_{r(c),t} + \psi_{i(c),t} + \tau_1 \text{Log GDP\_pw}_{c,t} \\ &\quad + \tau_2 \Big(\text{(Std) Pref. Safe Job}_c \times \text{Log GDP\_pw}_{c,t} \Big) + e_{c,t}, \end{aligned} \tag{6.9}$$

where $\text{Inf}_{c,t}$ is the aggregate informality rate, and μ_c are country fixed effects, $\delta_{r(c),t}$ are region-by-year fixed effects, $\psi_{i(c),t}$ are income-group-by-year fixed effects, $\text{(Std) Pref. Safe Job}_c$ is the standardized WVS proxy for country c, and $e_{c,t}$ is an error term.

In this specification, τ_1 captures the baseline semi-elasticity of informality with respect to productivity for a country with average job-security preferences, while τ_2 captures how this elasticity varies with the proxy: a one–standard-deviation increase in $\text{(Std) Pref. Safe Job}_c$ shifts the productivity–informality gradient by τ_2. Thus, productivity improvements are allowed to have stronger effects on informality in countries where workers place a higher value on job security and, by extension, on the attributes associated with formal employment.

Table 5 reports the results. Consistent with the model, the interaction between productivity and the proxy for the private value of formality is negative for all outcomes and statistically significant for aggregate informality and for industry. For a country with average preferences for job security ($\text{(Std) Pref. Safe Job}_c = 0$), the semi-elasticity of aggregate informality with respect to log GDP per worker is -0.205 (Column 1). For a country one standard deviation above the mean in job-security preferences, the implied semi-elasticity is -0.510 ($\tau_1 + \tau_2$), more than twice as large in absolute value and statistically different from zero. In other words, productivity improvements are associated with substantially larger reductions in informality in countries where workers place greater value on job security.

The sectoral results point in the same direction. In industry (Column 3), the baseline semi-elasticity is -0.227, which steepens to -0.364 in countries with stronger preferences for job security; the interaction term is again negative and significant. By contrast, in agriculture the baseline coefficient is positive but the interaction is negative and of similar magnitude, so that the implied semi-elasticity for high–job-security countries is close to zero and statistically insignificant. This is consistent with the idea that, given very low agricultural productivity, even sizable aggregate productivity gains do not move many agricultural workers across the formal–informal margin, regardless of their valuation of formality. Estimates for services are negative but imprecise.

Table 5: Informality and Preferences for Job Security

	Informality Rate			
	All Sectors (1)	Agriculture (2)	Industry (3)	Services (4)
(1) Log GDP per Worker	-0.205*	0.328**	-0.227**	-0.181
	(0.100)	(0.116)	(0.089)	(0.164)
(2) Log GDP per Worker X (Std) Pref. Safe Job	-0.306**	-0.280***	-0.137**	-0.056
	(0.102)	(0.060)	(0.063)	(0.064)
Sum Coefs. (1)+(2)	-0.510***	0.048	-0.364***	-0.237
P-Val.	[0.005]	[0.764]	[0.008]	[0.201]
Observations	194	194	194	194
Countries	23	23	23	23
Mean of Dep. Var. (%)	46.8	73.4	40.3	34.8
Mean Pref. Safe Job (%)	36.5	36.5	36.5	36.5

Notes: Each column reports coefficients from regressions of informality rates on log GDP per worker and its interaction with a proxy for preferences for job security. The dependent variable is the informality rate (share of informal workers in sectoral employment) for all sectors, agriculture, industry, or services. *(Std) Pref. Safe Job* is a country-level measure constructed from the World Values Survey as the average share of respondents who report "a safe job with no risk of closure or unemployment" as their top job attribute, standardized to have mean zero and unit standard deviation across countries. The interaction term in row (2) is the product of this standardized measure and log GDP per worker. "Sum Coefs. (1)+(2)" reports the implied semi-elasticity of informality with respect to log GDP per worker for a country with a one–standard-deviation higher preference for job security, with the associated p-value in brackets. All specifications include country fixed effects and year fixed effects interacted with region and World Bank income group. Standard errors are clustered at the country level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

These patterns are in line with the "informality-by-choice" mechanism. As productivity rises, fewer workers are excluded by the minimum-wage–productivity threshold, but whether this translates into large reductions in informality depends on the private value of formal jobs. The evidence suggests that in countries where workers place a higher priority on job security, productivity gains are associated with markedly stronger declines in aggregate and industrial informality, whereas in settings where formality is valued less, the same gains have a much weaker impact.

7 Discussion and Conclusion

This paper provides a unified way to interpret cross-country and within-country variation in labor informality. The framework emphasizes three forces: the legal wedge created by minimum wages and non-wage costs, the composition and productivity of sectors where workers and firms sort, and the private value of formality that makes compliance attractive or dispensable. A parsimonious model delivers testable predictions—stronger effects of the legal package and enforcement where thin-margin exposure is high, and lower informality where the perceived value of formality is greater—and the evidence supports them, especially within services.

The broader message is that formalization is not governed by a single instrument. Where many jobs operate near the legal floor, enforcement and legal design are complements; where most jobs lie comfortably above the floor, the same enforcement effort or legal reform will have limited bite, and sectoral development and reallocation may dominate aggregate dynamics. The results also underscore that raising the private value of formality—through low-friction administration, portability of benefits, and credible contract enforcement—can reduce informality directly and change how workers and firms respond to the legal wedge.

These findings have several policy implications. First, enforcement resources are likely to be most effective when targeted toward thin-margin activities, where compliance choices are most sensitive to monitoring and sanctions. Second, reforms that reduce the wedge (or that redesign it to better align labor costs and take-home pay) will matter more when the wedge binds for a sizable share of employment—and can otherwise generate limited formalization while still affecting labor costs. Third, a credible and portable benefits package can shift behavior even when enforcement is imperfect, because it changes the private calculus of compliance. Finally, because services are internally heterogeneous, policies that treat services as a single category risk misdiagnosing where formalization is feasible and where it is not.

Several priorities for future research follow directly from the framework. First, better *measurement* of enforcement and of the private value of formality at scale would tighten identification and comparability: administrative frictions, access costs, and benefit portability are rarely observed consistently across countries and time. Linking household and firm surveys to administrative records on registration, contributions, and inspections would allow sharper tests of the exclusion and choice margins in the model, including the wage-distribution "thin-margin" predictions. Second, progress on the *taxonomy of services* is essential. Services are internally heterogeneous; mapping which segments are thin-margin and which are formal-leaning—and how that map evolves with technology, platform intermediation, and verification—would improve predictions for sectoral reallocation and for the effectiveness of enforcement and valuation reforms. Third, the dynamics of *reallocation versus selection* merit closer study. Event-study designs around legal changes or administrative simplifications can be paired with structural decompositions to quantify how much of aggregate formalization comes from within-sector compliance versus shifts in employment toward activities where formality is commercially viable.

A fourth avenue is to bring *contract enforcement and third-party reporting* more explicitly into models of firm and worker behavior. In many service activities, verifiability and repeated interactions mediate compliance decisions as much as posted legal costs do; modeling those frictions would help reconcile heterogeneous responses to seemingly similar legal packages. Finally, distributional and welfare consequences—who gains and who loses along a feasible path of formalization—deserve sustained attention. The insurance value of informality during downturns suggests that sequencing matters: policy should avoid eliminating the buffer role of informality in bad times while building institutions that nudge matches toward formality in recoveries.

The promise of this approach is not a universal recipe but a disciplined map. By aligning concepts with measurement, and by tying stylized facts to a model that makes the operative margins explicit, the analysis identifies where leverage is empirically largest and where further data and theory can refine guidance. We hope this motivates a research agenda that is both comparative and mechanism-oriented, and that brings the study of informality closer to the level of operational detail required for country-specific policy design.

References

Adams, Zachary, Simon Deakin, Colm McLaughlin, Christopher Markou, Christopher Barnard, Anna Christensen, Aristea Koukiadaki, et al. 2023. *CBR Labour Regulation Index (CBR-LRI) Dataset, 1970–2022.* Technical report. Version 3.0, released December 2023. Centre for Business Research, University of Cambridge. https://doi.org/10.17863/CAM.101370.

Alfaro, Laura, Oscar Becerra, and Marcela Eslava. 2020. *EMEs and COVID-19: Shutting Down in a World of Informal and Tiny Firms.* Technical report w27360. National Bureau of Economic Research, June. https://doi.org/10.3386/w27360.

Allen, Jeffrey, Shanthi Nataraj, and Tyler C. Schipper. 2018. "Strict duality and overlapping productivity distributions between formal and informal firms." *Journal of Development Economics* 135 (November): 534–554. https://doi.org/10.1016/j.jdeveco.2018.08.011.

Almeida, Rita, and Pedro Carneiro. 2012. "Enforcement of Labor Regulation and Informality." *American Economic Journal: Applied Economics* 4, no. 3 (July): 64–89. https://doi.org/10.1257/app.4.3.64.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104:184–198. https://doi.org/10.1016/j.jdeveco.2012.10.001.

Bernal, Raquel, Marcela Eslava, Marcela Meléndez, and Alvaro Pinzón. 2017. "Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform." *Economia* 18 (1): 41–74. https://doi.org/10.18235/0009369.

Botero, Juan C., Simeon Djankov, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "The Regulation of Labor*." *The Quarterly Journal of Economics* 119, no. 4 (November): 1339–1382. https://doi.org/10.1162/0033553042476215.

Dell'Anno, Roberto. 2022. "Theories and definitions of the informal economy: A survey." *Journal of Economic Surveys* 36, no. 5 (December): 1610–1643. https://doi.org/10.1111/joes.12487.

Dix-Carneiro, Rafael, Pinelopi Goldberg, Costas Meghir, and Gabriel Ulyssea. 2021. *Trade and Informality in the Presence of Labor Market Frictions and Regulations.* Technical report. Cambridge, MA: National Bureau of Economic Research, January. https://doi.org/10.3386/w28391.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman. 2023. "Labor Market Dynamics and Development." *The Quarterly Journal of Economics* (May). https://doi.org/10.1093/qje/qjad019.

Elgin, Ceyhun, M. Ayhan Kose, Franziska Ohnsorge, and Shu Yu. 2021a. "Understanding Informality." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.3914265.

———. 2021b. *Understanding Informality.* CEPR Discussion Paper 16497. Available at: https://cepr.org/publications/dp16497. London: Centre for Economic Policy Research (CEPR).

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2004. "The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from Colombia." *Journal of Development Economics* 75, no. 2 (December): 333–371. https://doi.org/10.1016/j.jdeveco.2004.06.002.

Eslava, Marcela, Marcela Melendez, Laura Tenjo, and Nicolás Urdaneta. 2023. *Business Size, Development, and Inequality in Latin America: A Tale of One Tail.* Policy Research Working Paper Series 10584. The World Bank, October. https://doi.org/10.1596/1813-9450-10584.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer. 2015. "The next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–3182. https://doi.org/10.1257/aer.20130954.

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110, no. 2 (April): 458–474. https://doi.org/10.1086/338747.

Gollin, Douglas, and Joseph P. Kaboski. 2023. "New Views of Structural Transformation: Insights from Recent Literature." *Oxford Development Studies* 51, no. 4 (October): 339–361. https://doi.org/10.1080/13600818.2023.2280748.

Haanwinckel, Daniel, and Rodrigo R Soares. 2021. "Workforce Composition, Productivity, and Labour Regulations in a Compensating Differentials Theory of Informality." *The Review of Economic Studies* 88, no. 6 (November): 2970–3010. https://doi.org/10.1093/restud/rdab017.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114 (1): 83–116. https://doi.org/10.1162/003355399555954.

Haltiwanger, John, and Marcela Eslava. 2017. "The Drivers of Life-cycle Growth of Manufacturing Plants." *2017 Meeting Papers.*

Heckman, James, Carmen Pagés-Serra, Alejandra Cox Edwards, and Pablo Guidotti. 2000. "The Cost of Job Security Regulation: Evidence from Latin American Labor Markets." *Economía* 1 (1): 109–154. https://doi.org/10.1353/eco.2000.0007.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi. 2014. "Chapter 6 - Growth and Structural Transformation." In *Handbook of Economic Growth,* edited by Philippe Aghion and Steven N. Durlauf, 2:855–941. Handbook of Economic Growth. Elsevier. https://doi.org/10.1016/B978-0-444-53540-5.00006-9.

ICLS, ed. 1993. *Fifteenth International Conference of Labour Statisticians, Geneva, 19–28 January 1993: Report I — General Report.* ICLS/15/1. Geneva: International Labour Office.

———, ed. 2003. *Seventeenth International Conference of Labour Statisticians, Geneva, 24 November–3 December 2003: Report of the Conference.* ICLS/17/2003/4. Geneva: International Labour Office.

Kugler, Adriana, Maurice Kugler, and Luis Omar Herrera Prada. 2017. *Do Payroll Tax Breaks Stimulate Formality? Evidence from Colombia's Reform.* Technical report. Cambridge, MA: National Bureau of Economic Research, April. https://doi.org/10.3386/w23308.

La Porta, Rafael, and Andrei Shleifer. 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity* 2008 (2): 275–352. https://doi.org/10.1353/eca.0.0018.

———. 2014. "Informality and Development." *Journal of Economic Perspectives* 28, no. 3 (August): 109–126. https://doi.org/10.1257/jep.28.3.109.

Levy, Santiago. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico.* Brookings Institution Press. https://doi.org/10.7864/j.ctt6wpfgq.

Maloney, William F. 2004. "Informality Revisited." *World Development* 32, no. 7 (July): 1159–1178. https://doi.org/10.1016/j.worlddev.2004.01.008.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review* 105, no. 4 (April): 1509–1546. https://doi.org/10.1257/aer.20121110.

Mondragon-Velez, Camilo, Ximena Pena, and Daniel Santiago Wills. 2010. "Labor Market Rigidities and Informality in Colombia." *Economía (LACEA Journal)* 11 (1): 1–31. https://doi.org/10.1353/eco.2010.0009.

Paula, Áureo de, and José A. Scheinkman. 2010. "Value-Added Taxes, Chain Effects, and Informality." *American Economic Journal: Macroeconomics* 2, no. 4 (October): 195–221. https://doi.org/10.1257/mac.2.4.195.

Pérez Pérez, Jorge. 2020. "The minimum wage in formal and informal sectors: Evidence from an inflation shock." *World Development* 133 (September): 104999. https://doi.org/10.1016/j.worlddev.2020.104999.

Perry, Guillermo E., Omar Arias, Pablo Fajnzylber, William F. Maloney, Andrew Mason, and Jaime Saavedra-Chanduvi. 2007. *Informality: Exit or Exclusion.* The World Bank, May. https://doi.org/10.1596/978-0-8213-7092-6.

Rauch, James E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35 (1): 115–138. https://doi.org/10.1016/0304-3878(91)90065-4.

Rodrik, Dani. 2016. "Premature deindustrialization." *Journal of Economic Growth* 21, no. 1 (March): 1–33. https://doi.org/10.1007/s10887-015-9122-3.

Summers, Lawrence H. 1989. "Some Simple Economics of Mandated Benefits." *The American Economic Review* 79 (2): 177–183.

Tokman, Victor E. 1978. "An exploration into the nature of informal—formal sector relationships." *World Development* 6, nos. 9-10 (September): 1065–1075. https://doi.org/10.1016/0305-750X(78)90063-3.

Tybout, James R. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38, no. 1 (March): 11–44. https://doi.org/10.1257/jel.38.1.11.

Ulyssea, Gabriel. 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review* 108, no. 8 (August): 2015–2047. https://doi.org/10.1257/aer.20141745.

———. 2020. "Informality: Causes and Consequences for Development." *Annual Review of Economics* 12, no. 1 (August): 525–546. https://doi.org/10.1146/annurev-economics-082119-121914.

A Appendix

A.1 Data sources and variable construction

This section provides a detailed description of the datasets used for the graphical analysis and empirical findings presented in the article. The datasets come from a variety of publicly available sources and cover a wide range of countries and time periods. Below is a description of the variables created, organized by source.

A.1.1 ILOSTAT

Informal employment. Data on informal employment come from the ILO SDG labor Market Indicators database (ILOSDG), corresponding to SDG indicator 8.3.1: "*Proportion of informal employment in total employment by sex and sector.*" The indicator is reported annually and measures the share of workers whose main job does not provide formal coverage or protection under national labor legislation, taxation, or social protection systems.

Informal employment includes all jobs that lack basic social or legal protections or employment benefits. Following the ILO definition, this category encompasses:

1. *Wage employment in informal jobs*, i.e., employees without a written contract, social security coverage, or entitlement to paid leave or other employment benefits.
2. *Employment in formal enterprises in an informal capacity*, such as workers engaged without a formal contract or outside the scope of labor regulation.
3. *Self-employment in informal enterprises*, i.e., own-account workers and employers operating in unregistered or small household enterprises.

The ILO compiles this information from national labor force surveys using harmonized definitions based on the 15th and 17th International Conferences of Labor Statisticians (ICLS 1993, 2003). The resulting indicator represents the proportion of total employment classified as informal in each country-year, overall and by sector.

Sectors are aggregated into three groups:

- *Agriculture*: crop and animal production, forestry, fishing and aquaculture, and (where applicable) hunting-related activities.
- *Industry*: mining and quarrying, manufacturing, construction, and utilities (electricity, gas, steam and air conditioning; water supply, sewerage, and waste management).
- *Services*: all remaining non-agricultural, non-industrial activities, including wholesale and retail trade; transport and storage; accommodation and food

services; information and communications; financial and insurance services; real estate; professional and administrative services; education; health; arts, recreation, and other services; and public administration.

Statutory nominal gross monthly minimum wage. We use the statutory nominal gross monthly minimum wage compiled by ILOSTAT. This variable records, for each country-year, the legally mandated minimum remuneration payable to workers for a full month of work in the formal sector, before deductions and excluding irregular or non-wage payments (such as bonuses, overtime premia, or in-kind benefits). In countries with multiple minimum wages by sector, occupation, region, or age, ILOSTAT reports the statutory national rate or, where relevant, the rate that applies to the largest group of workers, in accordance with its standard metadata. Values are converted into annual terms and expressed in USD 2021 prices.

A.1.2 World Bank's Informal Economy Database

Self-employment. The indicator on self-employment is taken directly from the World Bank's *Informal Economy Database* (Elgin et al. 2021b), which provides annual indicators for up to 191 economies over 1990–2020 on various aspects of informality and related labor–market constructs. Self-employment refers to individuals whose principal economic activity is carried out on their own account or in collaboration with family members, rather than for an employer. Following the ILO's *Status in Employment* classification, this category includes:

1. **Employers** (self-employed persons with employees),
2. **Own-account workers** (self-employed persons without employees),
3. **Contributing family workers** (unpaid workers assisting in a family enterprise), and
4. **Members of producers' cooperatives**.

A.1.3 Penn World Tables (PWT)

We use the Penn World Table (PWT) version 10.01, a country-year panel covering 183 countries from 1950 to 2019 (Feenstra, Inklaar, and Timmer 2015). The variables used are discussed below.

GDP per worker. GDP per worker is defined as real GDP at 2017 prices (in millions of USD) divided by total employment (in millions). Total employment corresponds to the number of employed individuals—both employees and self-employed—who contribute to the production of goods and services as defined in the System of National Accounts (SNA).

Average earnings per worker. We compute average earnings per worker as

$$\text{Avg. earnings per worker}_{ct} = \frac{\text{Labor compensation}_{ct}}{\text{Total employment}_{ct}}.$$

In PWT, *labor compensation* is defined as the sum of compensation of employees and an imputed labor component of the income of the self-employed. Because the labor income of self-employed workers is not directly observable, PWT follows the adjustment methods in Gollin (2002) as implemented by Feenstra, Inklaar, and Timmer (2015): when national accounts report mixed income, it is split between labor and capital in the same proportion as in the rest of the economy; when mixed-income data are unavailable, PWT combines a "same-wage" correction (assuming self-employed workers earn the same average wage as employees) with an agriculture-based correction (adding agricultural value added to employee compensation) and uses the minimum of these two adjustments to construct total labor compensation.

Consistent with SNA 2008, labor compensation includes the full cost to employers of employing labor, namely:

1. **Wages and salaries** in cash or in kind, and
2. **Employers' social contributions**, including payments to social security schemes, pension funds, health insurance, and other legally or contractually required benefits.

Human capital index. The variable measures the *index of human capital per person* and captures differences in the quality of labor input across countries and over time based on educational attainment. Following Hall and Jones (1999) and using years of schooling from Barro and Lee (2013), PWT maps average schooling into a productivity-adjusted index with *piecewise* Mincer-type returns.

Let S_{ct} be average years of schooling. Define segment lengths

$$s_1 = \min\{4, S_{ct}\}, \quad s_2 = \min\{4, \max(S_{ct} - 4, 0)\}, \quad s_3 = \max(S_{ct} - 8, 0),$$

and segment returns $r_1 = 0.134$, $r_2 = 0.101$, $r_3 = 0.068$. The index is

$$\text{Human capital index}_{ct} = \exp\big(r_1 s_1 + r_2 s_2 + r_3 s_3\big).$$

A.1.4 World Bank's Global Productivity (ASPD) Database

Output per worker by sector. We measure sectoral labor productivity using the ASPD dataset, defined as real value added per worker. Formally, for country c, sector s, and year t:

$$\text{Labor productivity}_{cst} = \frac{\text{Real value added}_{cst}}{\text{Employment}_{cst}},$$

where real value added is expressed in 2017 constant USD (in thousands) and employment is the number of persons employed in sector s. The ASPD sectoral dataset reports annual statistics for nine broad ISIC-based sectors:

1. *Agriculture* (agriculture, forestry, fishing),
2. *Mining* (mining and quarrying),
3. *Manufacturing*,
4. *Utilities* (electricity, gas, steam, air conditioning),
5. *Construction*,
6. *Trade services* (wholesale/retail trade; repair of motor vehicles and motorcycles; accommodation and food services),
7. *Transport services* (transportation and storage; information and communication),
8. *Financial and business services* (financial and insurance; real estate; professional, scientific, and technical; administrative and support),
9. *Other services* (public administration and defense; education; health; arts and recreation; other services; household production for own use; extraterritorial bodies).

We aggregate these nine sectors into three groups consistent with our informality data: *agriculture* (1), *industry* (2–5), and *services* (6–9) using employment-weighted averages.

Coverage spans up to 103 countries through 2017 (some to 2018), depending on sector- and country-specific availability. ASPD combines national accounts and labor statistics from WDI, OECD STAN/KLEMS, GGDC, EASD, ILOSTAT, and national sources to construct consistent sectoral series.

A.1.5 CBR Labour Regulation Index (CBR-LRI)

Labor regulation index. We use information from the *CBR Labor Regulation Index* (CBR-LRI), developed by the Centre for Business Research at the University of Cambridge (Adams et al. 2023). The CBR-LRI is a longitudinal dataset that codifies the content of national labor laws across 117 countries from 1970 to 2022. It applies a "leximetric" methodology, converting qualitative legal rules into quantitative indicators on a scale from 0 (no protection) to 1 (maximum protection). Each indicator captures the formal, de jure degree of legal protection afforded to workers in different domains of labor regulation, including employment forms, working time, dismissal protection, collective representation, and collective action. The coding is based on statutory law and case law, and in some cases administrative regulations and collective agreements that have binding legal force. The resulting database records the intended normative content of labor legislation, not its enforcement or practical impact.

We construct an index of *non-wage labor costs* by aggregating a subset of 13 indicators from the CBR-LRI that capture legally mandated benefits and procedural constraints on employment relations. Specifically, the index includes:

1. *Dismissal costs for part-time workers* (CBR3);
2. *Entitlements to paid time off*, including:
    - Annual leave (CBR9),
    - Public holidays (CBR10),
    - Overtime pay (CBR11), and
    - Weekend pay (CBR12);
3. *Legally mandated redundancy compensation* (CBR17);
4. *Procedural and substantive constraints on dismissal* (CBR19, CBR20);
5. *Rights to unionization and collective bargaining*, including the duty to bargain and the extension of collective agreements (CBR25–CBR28); and
6. *Rights to, and restrictions on, industrial action*, including unofficial action and the right to strike (CBR32, CBR36).

Each indicator is coded between 0 and 1, with higher values denoting stronger legal protection.

We compute the unweighted sum of these components to obtain a synthetic measure of the legal generosity of employment-related benefits and protections that impose direct or indirect costs on employers. To improve comparability and interpretation, we standardize the resulting index by converting it into z-scores, so that a one-unit change corresponds to a one–standard deviation difference in the underlying distribution.