International Development

Prizewinning Dissertation 2025

No.25-LG

The Role of Social Security in Labour Market Formalization: Evidence from Brazil's MEI Programme

Ludovica Gerbino

Published: Feb 2026

Department of International Development

London School of Economics and Political Science

Houghton Street **Tel: +44 (020) 7955 7425/6252**

London **Fax: +44 (020) 7955-6844**

WC2A 2AE UK **Email: d.daley@lse.ac.uk**

Website: http://www.lse.ac.uk/internationalDevelopment/home.aspx

Abstract

This research examines the role of contributory social security in promoting labour market formalization, using evidence from Brazil's *Microempreendedor Individual* (MEI) programme. Exploiting a 2011 reform that halved mandatory social security contributions, I estimate difference-in-differences and triple-difference models to estimate average and heterogeneous effects of Brazil's formalization outcomes from 2007–2015. The results show that lowering contribution costs significantly increased the number and share of formal micro-entrepreneurs, largely through the entry of new formal businesses rather than transitions from informality. Contrary to expectations, women of childbearing age - eligible for maternity benefits - formalized less than other groups, suggesting that income constraints and household responsibilities outweighed potential incentives. No strong evidence is found for a retirement channel. The findings highlight that while contributory schemes can encourage formality, their effectiveness depends on affordability and structural barriers, underscoring the need for differentiated contribution schedules and complementary policies to ensure inclusiveness.

Contents

1 Introduction

Informality remains a persistent and widespread feature of labour markets in low- and middle-income countries (LMICs), particularly in Latin America, where over half of the labour force operates in the informal economy[1](ILO 2023). This is especially prevalent among micro-entrepreneurs and self-employed workers, over 75% of whom are engaged in informal employment (ILO, 2024). High levels of informality pose major policy challenges: they restrict a country's tax base and fiscal capacity, distort competition, and contribute to inefficient resource allocation (Besley and Persson, 2013; Hsieh and Klenow, 2009). High informality is also closely linked to limited access to social protection[2], as informal workers are typically excluded from contributory social insurance schemes, resulting in precarious working conditions, heightened vulnerability to shocks, and income insecurity (ILO, 2018; La Porta and Shleifer, 2014). Moreover, small informal firms often operate with low levels of productivity, limited access to credit, and short-term business horizons, further reducing their capacity to contribute to formal systems (Maloney, 2004).

While the correlation between informality and lack of social protection is generally confirmed across the literature (Canelas and Niño-Zarazúa, 2022; Torm and Oehme, 2024), there is limited empirical consensus on whether and how social security can impact the conditions and incentives that underpin informal labour markets. On the one hand, extending access to contributory social security may increase the perceived value of formalization, especially for workers who place high utility on such benefits. On the other hand, if the cost of participating in contributory schemes is too high, or if informal workers can access non-contributory benefits without formal registration or their informality network can act as a buffer against shocks, social protection policies may not be perceived as a valuable incentive to formalize. These competing channels underscore the need to understand the role of social insurance schemes as a policy lever for formalization, in order to design social policies that not only protect, but also incentivize inclusion into formal labour markets.

This paper aims to examine the role of social security as a potential mechanism driving labour market formalization, by focusing on heterogeneous responses of micro-entrepreneurs to a reduction in social security contribution costs. It investigates whether groups more likely to place higher value on social insurance formalize at higher rates, thus shedding light on heterogeneous incentive effects.

For this purpose, I study Brazil's *Microempreendedor Individual* (Individual Micro-Entrepreneur, henceforth MEI) programme. Brazil offers a valuable context for this analysis, with informality affecting 71% of business owners, reaching 81% for own-account workers (OECD, 2020). Introduced in 2009, MEI sought to encourage formalization among informal micro-entrepreneurs by

[1] According to Recommendation No. 204 (ILO, 2015) informal economy "refers to all economic activities by workers and economic units that are – in law or in practice – not covered or insufficiently covered by formal arrangements". This includes own-account workers, employers, cooperative members, contributing family workers, and wage earners in informal employment – whether they work for formal or informal enterprises, are subcontracted, part of supply chains, or employed as domestic workers – whose employment relationships are not formally recognized or regulated.

[2] Social protection, or social security, is defined as the set of policies and programmes designed to reduce and prevent poverty and vulnerability across the life cycle. It includes child and family benefits, maternity protection, unemployment support, employment injury benefits, sickness benefits, health protection, old-age benefits, disability benefits and survivors' benefits. Social protection systems address all these policy areas by a mix of contributory schemes financed through mandatory contributions from workers and employers (social insurance) and non-contributory publicly funded social assistance (ILO, 2017).

reducing administrative and fiscal burdens and granting access to Brazil's social security system (*Instituto Nacional do Seguro Social*, INSS) upon registration. Under the MEI framework, federal and local taxes were fixed at negligible values, leaving as the only fiscal cost borne by registrants the monthly social security contribution, initially set at 11% of the minimum wage. This amount represented a substantial burden for low-income entrepreneurs, making the policy attractive mostly to higher earners (Rocha et al., 2018). In April 2011, a major reform halved this contribution to 5% of the minimum wage, significantly lowering the cost of accessing social protection. This change occurred at a point when registration procedures had already been fully simplified and entry costs eliminated, thereby allowing a cleaner identification of the effect of the cost of accessing social protection on formalization decisions.

This analysis leverages the April 2011 tax reform to examine how reducing the cost of social insurance - and thereby increasing access to social protection - affects formalization rates.Specifically, it assesses whether the policy increased the number of formal micro-entrepreneurs (extensive margin) and the share of formal micro-entrepreneurs within sectors, indicating a compositional shift from informal to formal status rather than a mere overall expansion in the number of micro-entrepreneurs.

My identification strategy relies on the tax reduction as a quasi-natural experiment to estimate first, a difference-in-difference model (DiD) to evaluate average effects of this policy on formalization rates; and then a triple-difference model (DDD), assessing differential effects of the reform on formalization outcomes across demographic groups differentially likely to value the relevant social security benefits. I will focus on the incentive effects of maternity benefits, analysing formalization outcomes for women of child-bearing age who are more likely to value the paid maternity leave. As a secondary, exploratory dimension, I also consider the incentives of retirement benefits, focusing on mid-aged workers for whom beginning contributions could realistically lead to pension eligibility by the statutory retirement age.

The empirical analysis covers the six largest metropolitan areas in Brazil – Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Porto Alegre – from 2007 to 2015, providing sufficient coverage before and after the April 2011 cutoff. The research draws on the Pesquisa Mensal de Emprego (PME) dataset, a rich micro-level rotating monthly panel that follows households for a short-term period (up to 16 months) and captures both formal and informal labour force participation, allowing to track employment status and labour market transitions. However, due to dataset limitations, individual identities cannot be linked over time, which precludes longitudinal analysis at the individual level. The DiD and DDD models are thus applied to aggregate-level outcomes at the industry-by-metropolitan area-by-month level, to provide evidence on sectoral changes in the level and composition of formal micro-entrepreneurship. In addition, because the PME dataset does not report industry classification data at the required level of disaggregation to determine sectoral eligibility, it is complemented with data from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD), a nationally representative annual household survey that contains more granular industry information.

The results show that the 2011 MEI reform led to a significant increase in the number and share of formal micro-entrepreneurs, confirming that lowering contribution costs can expand formalization. However, the evidence suggests that much of this growth occurred at the extensive

margin, through the entry of new businesses registering directly as formal, rather than through widespread transitions of incumbents from informality. The heterogeneous analysis reveals more complex dynamics. Women of childbearing age, who were expected to respond most strongly due to eligibility for maternity benefits, instead formalized less than other groups. Exploratory evidence suggests two main constraints: an income burden, with poorer women less able to afford even the reduced contribution, and a care burden, with larger household responsibilities further limiting responsiveness. By contrast, no clear effects were found for workers nearing retirement, reflecting the limited salience of long-term pension benefits in contexts of income volatility and short planning horizons.

This paper contributes to different related strands of literature: the determinants of labour market formalization in LMICs, and its relationship with social protection, shedding light on a less explored mechanism: access to contributory social security.

The first strand has mainly focused on reducing the costs of formality. Numerous studies show that simplifying tax regimes, lowering registration costs, and streamlining compliance procedures can increase formalization, particularly among micro and small firms. In Mexico, Bruhn (2011) and Kaplan et al. (2011) find respectively that simplifying firm registration led to an increase in the number of registered businesses and modest shift of informal firms into the formal sector for those operating close to the formal threshold. Marchese (2021), in a cross-country analysis across Latin America suggests that preferential taxation can play a role in encouraging formalization, but only if part of a broader set of policies aimed at improving the overall business environment, productivity growth and rule enforcement.

In the Brazilian context, Fajnzylber et al. (2011) examine the SIMPLES regime and find that it increased formality among small firms, particularly those that were already more productive and closer to the formal–informal margin, while low take-up among less productive firms suggests the presence of additional barriers beyond taxes. More directly relevant, Rocha et al. (2018) show that MEI significantly increased formalization among micro-entrepreneurs, driven more by reduced fiscal costs rather than by administrative simplification. Pereda et al. (2022) find stronger effects for female entrepreneurs. Overall, while these findings have established that reducing the fiscal or administrative costs of formality can expand registration, they also underscore that reducing entry barriers alone does not guarantee sustained formalization – raising the question of what other mechanisms may be at play, and whether individuals respond to cost reductions alone or to the bundled benefits of formalization.

While widely acknowledged as one of the key advantages of formality (Barrientos, 2011; ILO, 2017), social insurance has received less empirical scrutiny as an incentive to formalize. In general, the literature confirms a strong correlation between high informality and low social protection coverage (Torm and Oehme, 2024). Although it is sometimes argued that the reason behind this is that informality partially acts as a buffer against shocks thus reducing demand for formal insurance, evidence has shown that social protection is even more essential in contexts where informal work is widespread and living conditions are at subsistence level (Gerard and Naritomi, 2021; Liepmann and Pignatti, 2021). This paper contributes to this second, smaller but growing body of literature that examines the relationship between social insurance and formalization decisions.

Most studies in the existing literature focus on how non-contributory social protection programmes may discourage formalization by providing benefits to informal workers. Levy (2008) offers one of the earliest and most influential arguments in this space, suggesting that Mexico's non-contributory health program distorted labour markets by reducing the incentive to enter formal employment. Empirical evidence supports this view: Aterido et al. (2011) and Bosch and Campos-Vazquez (2014) find that the Mexican Seguro Popular program significantly reduced formal employment, especially among low-income households, and similar effects are found in Colombia by Camacho et al. (2014). Together, these findings suggest that when social protection is accessible independently of formal employment status, workers may be less inclined to bear the costs of formalization, potentially supporting the interpretation that social protection is a valued benefit, which may increase the incentive to formalize if exclusively accessible through the formal sector. Evidence on contributory social protection as a pull factor for formalization remain more limited. Azuara and Marinescu (2013) find that expanding access to formal health insurance increased formal employment among older Mexican workers. Bérgolo and Cruces (2011) evaluate a Uruguayan reform that extended health coverage to the dependents of formal workers, finding that the policy led to increased formal labour supply particularly among individuals with dependents.

Against this backdrop, this paper makes three contributions. First, it isolates the role of social security contribution costs – rather than broader administrative reforms – using a quasi-natural experiment in Brazil. Second, it explores heterogeneous behavioural responses, focusing on workers likely to value specific benefits (namely, maternity and retirement). Third, by combining administrative and survey data, it offers a granular view of formalization dynamics and transitions across labour market statuses. In doing so, it provides novel evidence on whether contributory social insurance can act not only as a safety net, but as an incentive to formalize in LMICs.

The remainder of the paper is structured as follows. Section 2 provides background on the institutional context of Brazil's MEI programme. Section 3 presents the conceptual framework motivating the empirical analysis. Section 4 describes the data sources, variable construction, and sample restrictions. Section 5 details the empirical strategy, including the identification approach, model specifications along with identification assumptions and validity tests. Section 6 presents the main results, covering average and heterogeneous effects, as well as exploring potential channels underlying the observed results. It also reports robustness checks. Section 7 outlines the limitations of the study. Section 8 concludes.

2 Institutional Setting: the MEI Programme

The MEI policy was introduced in July 2009 as a targeted extension of Brazil's broader presumptive tax regime for small businesses, SIMPLES Nacional. SIMPLES, originally launched in 1996, aimed to reduce the formalization burden on micro and small enterprises by simplifying the tax structure and consolidating multiple federal, state, and municipal levies into a single monthly payment. Although the programme significantly reduced compliance costs and supported firm survival and growth (Fajnzylber et al., 2011; Monteiro and Assunção, 2012),

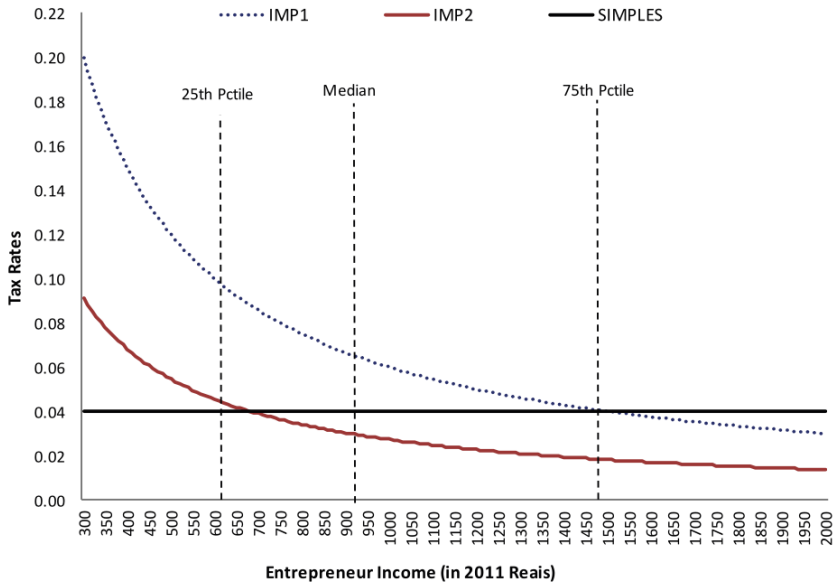
the regime primarily benefited relatively larger small firms with stable operations and employee structures, excluding large portions of the informal sector, particularly low-income and own-account workers. In response to these limitations, the federal government implemented the MEI policy in 2009 as a new institutional tier within the SIMPLES framework.

MEI was designed specifically for informal own-account workers with at maximum one employee, reducing barriers to formalization by streamlining administrative procedures but and decreasing the financial burden through simplified taxation. Importantly, MEI also explicitly linked formal economic registration to inclusion in Brazil's contributory social system (INSS). Social protection benefits granted upon registration include old-age retirement pensions (conditional on 180 months of contributions) and paid maternity leave (10 months), along with disability and sickness benefits, confinement aid, and survivor pensions. The policy was designed around three core eligibility criteria: (i) entrepreneur must be self-employed or have at most one employee; (ii) annual business revenue must remain below a fixed threshold (initially set at R$36,000 and further expanded to R$60,000 in 2011); and (iii) business activity must fall within a list of over 400 sectors identified at the 7-digit economic activity classification level (CNAE), which mirrored the previously defined sectoral admissibility within SIMPLES framework.

The policy unfolded in two distinct phases, each reflecting a different balance between access and cost. Phase I (2009-2011) focused primarily on reducing administrative and fiscal entry costs: registration procedures were simplified and digitalized, while all federal taxes were eliminated and state and municipal taxes at negligible levels. The key remaining fiscal obligation was a monthly social security contribution fixed at 11% of the prevailing minimum wage, applied uniformly across all eligible entrepreneurs regardless of income levels. This introduced *de facto* a regressive tax schedule (Figure 1). As a result, only entrepreneurs in the upper percentiles of the MEI-eligible income distribution actually experienced a tax reduction when compared to SIMPLES (Rocha et al., 2018). Phase I achieved significant success in removing bureaucratic hurdles to formalization, but failed to reduce the cost of formalization for the majority of low-income micro-entrepreneurs.

Phase II, introduced in April 2011, marked a major policy reform, as the required monthly rate for the INSS contribution was reduced from 11% to 5% of the minimum wage, significantly lowering the cost of accessing formal social protection. The fixed monthly tax now represented a smaller share of earnings for a broader segment of low-income entrepreneurs, with those above the 25th percentile of the earnings distribution gaining a net tax advantage compared to SIMPLES. By the time this reform occurred, entry procedures had already been simplified and were functioning at full scale, offering a relatively clean policy shock that isolates the effect of reduced social security contribution costs on the formalization decisions.

Figure 1: Tax rates under SIMPLES, MEI1, and MEI2 across the entrepreneur income distribution (2011 Reais).



Notes: The figure plots statutory tax rates for entrepreneurs under the SIMPLES regime and the two phases of the MEI programme (Phase I (IMP1), Phase 2 (IMP2), expressed as a share of monthly income in 2011 Reais. Vertical dashed lines indicate the 25th percentile, median, and 75th percentile of the income distribution among entrepreneurs with at most one employee in April 2011. Tax rates (TR) are $TR_{\text{MEI1}} = \frac{0.11\,MW}{Y_i}$ and $TR_{\text{MEI2}} = \frac{0.05\,MW}{Y_i}$.
Source: Rocha et al. (2018).

3 Conceptual Framework

A foundational premise of this analysis is that formalization is not solely determined by regulatory enforcement or entry costs, but also by the structure of incentives embedded in social protection systems. The decision to formalize reflects a trade-off between costs – such as tax contributions, regulatory burdens, and administrative procedures – and perceived benefits, which may include legal protection, access to credit markets, and crucially, access to social insurance. The informal status is maintained when the expected costs of formalization exceed the perceived benefits (Levy, 2008; Ulyssea, 2018). The net valuation of these costs and benefits may vary significantly across individuals, depending on their income level, demographic characteristics, and household circumstances.

To conceptualize this behavioural framework, consider this simplified model in which a micro-entrepreneur chooses between informal status and formal status under the MEI regime. The net benefit of each status can be expressed as follows:

$$NB^{\text{formal}} = B^{SS} - C^{SS} - C^{admin}$$
$$NB^{\text{informal}} = B^{\text{non-SS}} - C^{vuln}$$

where:

- B^{SS}: perceived benefit from contributory social security (e.g., maternity leave, pension accrual);
- C^{SS}: social security contribution (fixed monthly payment under MEI);
- C^{admin}: administrative/regulatory burden of formalization (e.g., registration and compliance);
- $B^{\text{non-SS}}$: benefits of remaining informal (e.g., flexibility, informal networks, universal services);
- C^{vuln}: implicit costs of vulnerability while informal (e.g., income volatility, lack of legal protection, exposure to shocks).

The entrepreneur formalizes if and only if

$$NB^{formal} \geq NB^{informal} \quad \Leftrightarrow \quad B^{SS} - C^{SS} - C^{admin} \geq B^{non-SS} - C^{vuln}$$

The 2011 MEI reform reduced C^{SS}, while C^{admin} already been minimized (in Phase I), thereby improving the benefit of formal status. However, this effect is not expected to be uniform, as the perceived value of formal social protection B^{SS} differs by demographic profile and proximity to specific entitlements. This reform therefore offers a quasi-experimental setting to isolate how lowering the cost of access to contributory social insurance affects formalization decisions. The conceptual framework supports the following testable predictions:

(i) Formalization rates will increase in MEI-eligible sectors following the 2011 reform (average effects).
(ii) The effect will be stronger among individuals within demographic groups associated with higher expected returns to social protection (heterogeneous effects by demographic group).

This research mainly tests one key incentive, analysing the response to this reform of women of childbearing age, who may anticipate maternity-related benefits. As a secondary, exploratory mechanism I also consider mid-aged workers nearing the retirement eligibility, who may place greater value on pension accrual, to study the pension benefit incentive. It is important to note that the effects of this policy are expected to be less strong, as retirement under MEI requires several years of contributions, making the benefits less immediate and salient especially for informal workers facing uncertainty and limited long-term planning.

Crucially, there is an interplay between the maternity and pension benefits, particularly for women, who face dual incentives. Nonetheless, while dual incentives strengthen the case for formalization, this group also faces counteracting pressures, as formal employment may reduce flexibility in working hours and increase obligations incompatible with caregiving responsibilities usually attributed to women (Goldin, 2014). Therefore, the ultimate impact on women's formalization choices is ambiguous, contingent upon the relative strength of competing incentives – enhanced social protection versus restricted flexibility in working arrangements. The conceptual framework acknowledges these complex and countervailing forces, highlighting an empirical question: whether the incentives stemming from social protection benefits sufficiently outweigh potential reductions in employment flexibility and other monetary and non-monetary costs related to it. The analysis will empirically test and quantify these competing effects,

elucidating the conditions under which formalization benefits outweigh associated costs and constraints for specific demographic groups.

Overall, the effect of the reform is expected to be heterogeneous, with the degree of responsiveness mediated by the perceived relevance and timing of social protection benefits. The above predictions will be tested in Section 5 and 6.

4 Data

4.1 Data Sources and Sample Construction

This study employs a rich micro-level dataset, the Pesquisa Mensal de Emprego (PME), a monthly rotating panel survey conducted by the Brazilian Institute of Geography and Statistics (IBGE). In the survey, each household is interviewed for 4 consecutive months, dropped for 8 months, and then re-interviewed for an additional 4 months, resulting in a total longitudinal dimension of up to 16 months per household. PME covers Brazil's six largest metropolitan regions – São Paulo, Rio de Janeiro, Belo Horizonte, Salvador, Recife, and Porto Alegre. Key variables in PME include: employment status (employed, unemployed, self-employed, informal/formal), contributions to INSS, industry of employment (classified at the two-digit CNAE sector codes), and demographic information (age, gender, education, household composition, geographic location).

Although the dataset presents a rotating panel structure, it is not possible to reliably link single individuals overtime due to changes in household composition and non-persistent individual codes, preventing to track the same individual across interview waves. Consequently, the dataset effectively behaves as a stacked cross-section of individuals observed at different points in time, and I therefore aggregate the data at the sector-by-metropolitan area-by-month level, which preserves the variation necessary for the empirical strategy while avoiding misclassification from attempted individual tracking. The analysis uses PME microdata from January 2007 to December 2015, or 108 months, allowing for enough time span before and after April 2011 cutoff (51 months pre-reform, 57 after). It focuses on the six metropolitan areas consistently covered by the PME, which are similarly represented, each accounting for between 15% and 18% of total observations (concerns on external validity raised by the limited territorial coverage of this study are addressed in Section 7).

I restrict the sample to self-employed individuals and employers with at most one employee, i.e., the defined microentrepreneurs targeted by the MEI policy. Formal entrepreneurs are defined as those who contribute to the social security system INSS, a commonly used indicator of formality in the literature. Moreover, as mentioned in Section 2, the program linked together social security and business registration, making contribution to INSS a good formalization measure in this context. I also restrict to individuals between 18 and 65 years old (working age) and exclude those in non-comparable sectors in the labour market, such as public sector and military, agriculture and extracting activities, international organizations and other extra-territorial institutions. For the main specification, I exclude the first quarter of 2011 (January-March) to mitigate potential anticipatory responses to the reform. Robustness checks in Section

6.4 will assess the sensitivity of results to this exclusion.

4.2 Intensity Variable Construction

One limitation of the PME dataset is that industry classification is reported at only the two-digit CNAE level, while MEI eligibility criteria are defined at a higher disaggregated level of economic activity (7-digit CNAE subclass codes). This mismatch prevents a direct classification of individual eligibility within the PME sample. To address this, I follow the strategy employed by Rocha et al. (2018), and incorporate a third data source: the *Pesquisa Nacional por Amostra de Domicílios* (PNAD), a nationally representative annual household survey that contains detailed industry classifications aligned with the level at which MEI eligibility is defined[3] Using the 2009[4] wave – prior to the MEI reform expansion – I compute a sector-level treatment intensity variable that captures the proportion of entrepreneurs within each 2-digit sector who operate in MEI-eligible sub-industries. Formally, for each 2-digit CNAE industry s, I calculate:

$$Intensity_s = \frac{\sum_{k \in s} \mathbb{I}[\text{Sub-industry } k \text{ is eligible}] \cdot N_{ks}}{\sum_{k \in s} N_{ks}} \tag{1}$$

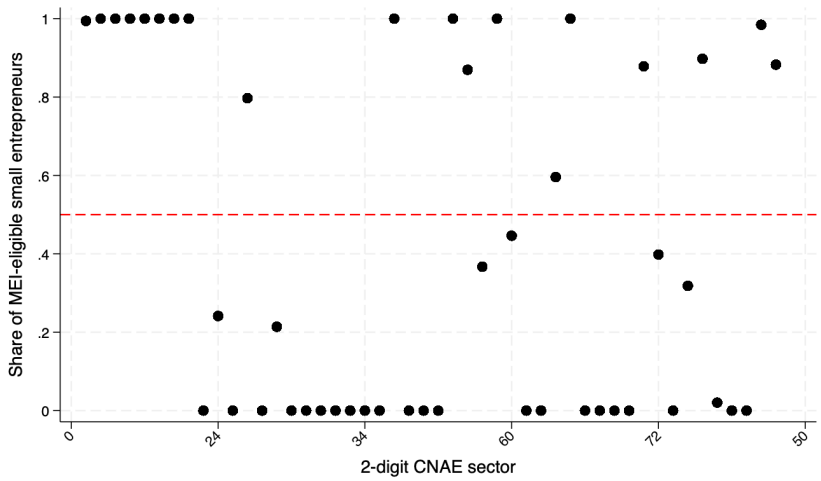
where N_{ks} denotes the number of self-employed individuals working in sub-industry k, which belongs to the broader 2-digit sector s. The indicator $\mathbb{I}[\text{Sub-industry } k \text{ is eligible}]$ equals 1 if sub-industry k is eligible for MEI registration, and 0 otherwise. This ratio reflects the pre-reform share of potentially eligible entrepreneurs within each broader sector. I then merge the resulting $Intensity_s$ values with the PME individual-level panel at the 2-digit sector level, assigning each worker the treatment intensity associated with their reported industry. This procedure allows to generate a continuous measure of exposure, to proxy for eligibility to the MEI reform in the absence of direct identifiers in PME.

The empirical strategy underlying the analysis using PME data thus relies on the assumption that individuals located in industries with higher potential of treatment are more likely to benefit from the programme than those located in industries with a lower potential. Figure 2 plots the distribution of treatment intensity values for each 2-digit CNAE sector, highlighting the polarization of intensity, with most sectors clustered either near zero or at the maximum value of 1, and relatively few in the intermediate range. This polarization makes the treatment intensity variable resemble a binary classification, with sectors falling clearly into low- or high-exposure categories. Table A.1 in the Appendix reports treatment intensity values for all 2-digit CNAE sectors included in the analysis, along with a sample of calculations made.

[3]The Brazilian National Bureau of Statistics (IBGE) adopts two distinct industry classification systems: one for the National Accounts, also used to determine program eligibility at the 7-digit level, and another for Household Surveys, including the PNAD, at the 5-digit level. These systems share the same level of disaggregation, and IBGE provides a direct mapping between them. I rely on this correspondence to match the PNAD industry classifications with those used to define industry eligibility under the MEI programme.

[4]I use the 2009 PNAD wave rather than 2010 because the survey was not conducted in 2010. Brazil conducted its full national Census in that same year and, as a rule, PNAD is not conducted in Census years to avoid duplication of effort and respondent burden.

Figure 2: Treatment Intensity by 2-Digit CNAE Sector



Note: This scatter plot displays the share of MEI-eligible small entrepreneurs (treatment intensity) for each 2-digit CNAE sector. The horizontal dashed red line at 0.5 serves as a reference threshold, visually separating sectors with high treatment exposure (above 0.5) from those with low exposure (below 0.5).
Source: Author's own calculations

4.3 Descriptive Statistics and Pre-treatment Covariate Balance

Table 1 reports descriptive statistics for the PME aggregated sample at baseline (Panel A) and separately by treatment status (Panel B). Although the intensity treatment variable is continuous in all analyses, for this table I define treatment using the cutoff *Intensity¿0.5*. This split yields roughly equal group sizes (49.72% in the control group, 50.38% in the treatment group). Pre-treatment covariates are balanced across groups, with the exception of formal INSS coverage, which differs in a way consistent with the programme's targeting. This does not threaten identification, as the empirical strategy accounts for baseline level differences and common shocks.

Table 1: Descriptive statistics at baseline (pre-treatment, 2007–2010)

	Panel A					Panel B	
Variable	**Obs.**	**Mean**	**SD**	**Min**	**Max**	**Control mean**	**Treatment mean**
Age (years)	10,261	42.390	6.469	18	65	42.636	42.128
Male (dummy = 1 if male)	10,261	0.653	0.328	0	1	0.740	0.561
Race or colour (dummy = 1 if black or brown)	10,261	0.466	0.341	0	1	0.441	0.493
Years of schooling	10,261	8.463	2.072	0	11	8.758	8.150
Log (usual monthly income)	10,261	6.605	0.698	2.398	10.596	6.815	6.381
Hours worked usually	10,261	42.149	7.560	4	98	42.911	41.337
Contributes to INSS (formal)	10,261	0.247	0.261	0	1	0.305	0.185

Notes: Panel A presents descriptive statistics at baseline for the PME sample of micro-entrepreneurs (individual self-employed or small employer with at most one employee). Panel B shows control and treatment means used for a pre-treatment balance check. All tabulations refer to pre-policy years (2007–2010).

5 Empirical Strategy

5.1 Identification Strategy

The aim of the estimates presented below is to identify the causal effect of the reduction in the cost of accessing social security benefits on levels of formal micro-entrepreneurship, and thereby test the two predictions derived from the conceptual framework (Section 3).

(i) Formalization rates increase in MEI-eligible sectors following the 2011 reform (average effect).

(ii) The effect is stronger among individuals within demographic groups associated with higher expected returns to social protection (heterogeneous effects by demographic group). The main focus is on women of childbearing age, who may value access to maternity benefits. As a secondary, exploratory dimension, I also consider workers nearing retirement.

The 2011 reform provides a quasi-experimental setting for causal inference. The policy reduced the cost of social security contributions for all MEI-registered micro-entrepreneurs from 11% to 5% of the minimum wage, without altering other program features such as eligibility criteria, benefit design, or administrative requirements. Because MEI eligibility was defined exogenously by pre-specified administrative rules based on industry classification (CNAE) and applied uniformly across the country, the reform introduced variation across sectors in exposure to the cost reduction. Sectors comprising a higher proportion of MEI-eligible activities experienced a larger relative reduction in the cost of formality, creating the basis for a treatment-versus-control comparison. To measure formalization rates, I construct two outcome variables: the log number of formal micro-entrepreneurs, and the share of formal micro-entrepreneurs within the total number of self-employed and small employers.

5.2 Empirical Specifications

5.2.1 Average Effects

The first set of estimates tests whether the reform increased formalization rates on average (Prediction i). I estimate a Difference-in-Differences (DiD) model comparing formalization trends before and after the 2011 reform in sectors with different degrees of exposure to MEI eligibility. Sectoral treatment intensity is defined as the share of micro-entrepreneurs in a sector who were eligible for MEI registration (Section 4). The analysis is conducted at the sector-by-metropolitan area-by-month level.

I estimate the following regression:

$$Y_{smt} = \alpha + \beta_1(Post_t \times Intensity_s) + \gamma_s + \delta_m + \lambda_t + X'_{smt}\theta + \varepsilon_{smt} \tag{2}$$

where Y_{smt} is the outcome variable (log number or share of formal micro-entrepreneurs) in each 2-digit sector s, metropolitan area m, and month t; $Post_t$ is a dummy that equals 1 for months after April 2011 reform; $Intensity_s$ is the treatment intensity proxy defined in equation (1).

I include *sector* fixed effects γ_s to control for all time-invariant sector characteristics (e.g., sector size, regulatory environments, historical formalization rates), *metropolitan area* fixed effects δ_m to absorb persistent differences between metropolitan regions (e.g., enforcement capacity and administrative reach), and *time* fixed effects λ_t capturing business cycle fluctuations and macroeconomic shocks that could jointly affect all sectors across metropolitan regions. Together, these fixed effects allow the identification to depend solely on the variation within the cell between the treated and comparison groups, net of broader sectoral and regional trends. The vector X_{smt} collects time-varying sector–metro covariates (gender, education, race, and age composition) to account for gradual compositional shifts. Standard errors are clustered at the sector × metro level. The coefficient of interest β_1 captures the change in the outcome of interest of formal micro-entrepreneurs in MEI-eligible sectors, after the 2011 reform, with respect to formalization trends in micro-entrepreneurs in non-eligible sectors (i.e., not affected by the policy).

5.2.2 Heterogeneous Effects by Demographic Group

This section tests whether the reform led to a differential change in formalization outcomes of interest (number and share of formal micro-entrepreneurs, defined above) among demographic groups expected to value differently social protection benefits (Prediction ii).

The study focuses on the incentive effects of maternity benefits, analysing formalization outcomes for women of child-bearing age who are more likely to value the paid maternity leave. I classify women aged 20–40 as of prime childbearing age, the treatment group, while the control group consists of women aged 40–65, classified as post-reproductive, and men of all ages, who serve as an additional comparison group not directly affected by maternity-benefit eligibility[5].

As a secondary, exploratory dimension, I also consider the incentives of retirement benefits, focusing on mid-aged workers for whom beginning contributions could realistically lead to pension eligibility by the statutory retirement age. In this case, the treated group comprises workers aged 45-55 – who are closer to retirement eligibility age and more likely to value INSS pension accrual –, compared to those at earlier or later career stages (workers aged 18-44 and 56-64).

I implement a Triple Difference (DDD) specification that interacts this baseline DiD (defined above) with demographic groups expected to place higher value on specific social insurance benefits.

The analysis is again conducted at the sector-by-metropolitan area-by-month level, ensuring comparability with average effect results. I estimate the following regression:

$$\begin{aligned} Y_{smtg} = \alpha &+ \beta_1(Post_t \times Intensity_s \times Group_g) + \beta_2(Post_t \times Intensity_s) \\ &+ \beta_3(Post_t \times Group_g) + \beta_4(Intensity_s \times Group_g) + \gamma_s + \delta_m + \lambda_t + X'_{smtg}\theta + \varepsilon_{smtg} \end{aligned} \tag{3}$$

where Y_{smtg} is the outcome of interest in each sector s, metropolitan area m, month t, and demographic group g; $Group_g$ denotes the demographic group of interest; and remaining

[5]Age-group definitions follow Gruber (1994) and are consistent with Brazilian fertility patterns during the study period.

terms are analogous to those in equation (2). The coefficient of interest, β_1, captures the triple difference effect, the differential post-reform change for group g in more-exposed sectors relative to less-exposed sectors and relative to other groups.

5.3 Identification Assumptions

The credibility of the identification strategy rests on several key assumptions required for a valid DiD and DDD estimation.

The core identifying assumption is that in the absence of the reform, formalization outcomes would have evolved similarly between (i) sectors with higher versus lower exposure to MEI eligibility and (ii) - within each sector type - demographic groups that differ in their expected returns to social protection (parallel trends assumption). As a result, any post-2011 divergence in outcomes between more- and less-exposed sectors, or between demographic groups within a sector type, can be attributed to the reform. To assess this, I examine pre-treatment trends by visually inspecting the coefficients on leads of the treatment in the event-study plots at two levels, DiD for sectoral outcomes (Figure 3 and 4), and heterogeneous DDD for demographic-specific effects (Figure 5 and 6). Across all panels, pre-treatment coefficients are close to zero and statistically insignificant, providing graphical support for parallel trends in both the sector-level and demographic-split analyses.

Figure 3: Event-study (DiD): Log number of formal micro-entrepreneurs

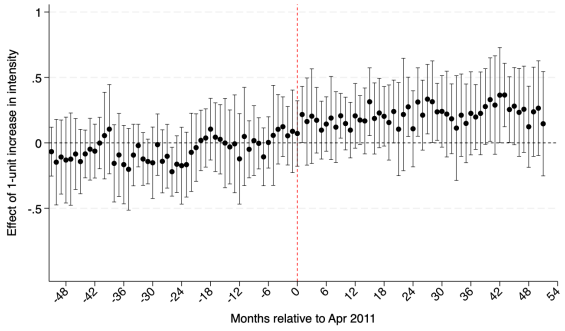


Figure 4: Event-study (DiD): Share of formal micro-entrepreneurs

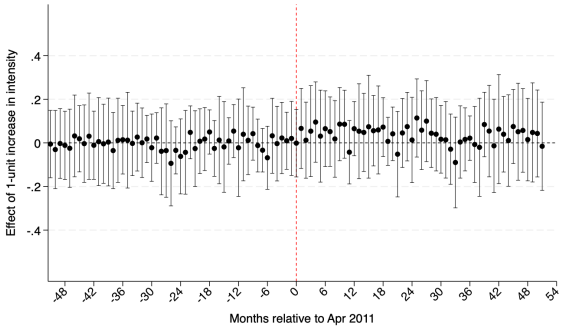


Figure 5: Event-study (DDD): Log number of formal micro-entrepreneurs - maternity benefits

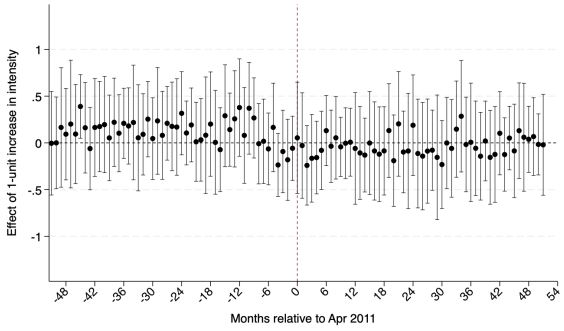
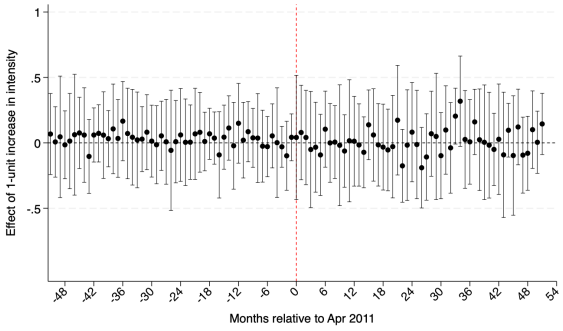


Figure 6: Event-study (DDD): Share of formal micro-entrepreneurs - maternity benefits



Notes: Event-study plots at the sector×metro×month level. Coefficients are monthly leads/lags relative to the month before April 2011. Fixed effects: sector, metro, month–year; composition controls included. Standard errors clustered at sector×metro; 95% CIs. To align with the main specification, the first quarter of 2011 is omitted to avoid anticipatory effects in the leads
Source: Author's calculations based on PME (2007–2015).

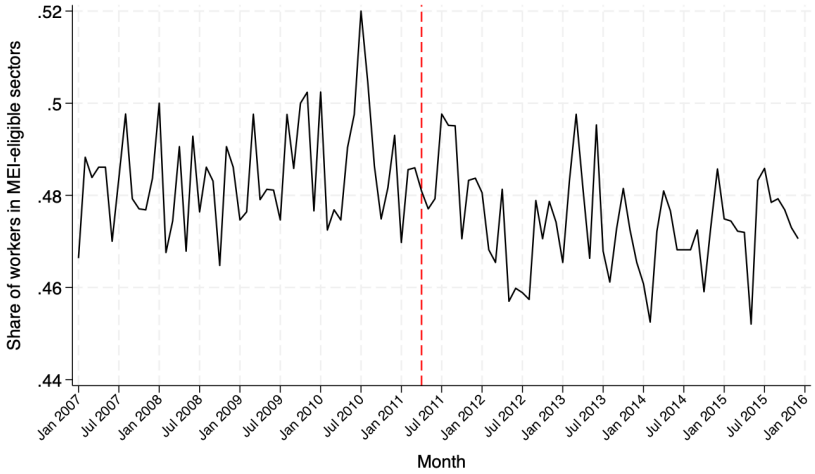
The second assumption is that assignment to treatment – defined by MEI sectoral eligibility and demographic subgroup – is exogenous to unobserved determinants of formalization outcomes (exogenous assignment to treatment). MEI eligibility was determined by detailed pre-existing CNAE classifications based on the SIMPLES list of eligible industries, identified more than a decade prior to MEI, and applied uniformly nationwide. Moreover, there is no indication that the government hand-picked specific industries based observable trends in formality or policy targeting (Rocha et al., 2018). Demographic groups (e.g., women of childbearing age) are determined by age and sex, which are not chosen in response to the 2011 reform. These features make assignment plausibly independent of unobserved shocks to sectoral formalization. To empirically support this claim, I conduct falsification tests using pre-reform periods only (2007-2010), assigning pseudo-reform dates to January 2008, January 2009, and January 2010. I find no significant coefficients for formalization outcomes. The Appendix contains all regression tables and results for such placebo tests (Table A.2 and A.3).

A further identifying assumption is that no other shock occurred simultaneously with the April 2011 tax reduction that differentially affected MEI-eligible sectors or specific demographic groups. There is no evidence of substantial economic fluctuation around MEI's implementation (Rocha et al., 2018). Furthermore, the empirical specifications include rich fixed effects - sector, metropolitan area, and month, - to absorb time-invariant sector and location heterogeneity as well as nationwide business-cycle movements common to all sectors and metros.

Finally, the reform should not induce systematic reclassification of workers or firms into MEI-eligible activities (or out of them) that would mechanically inflate treatment exposure. Identification assumes that the sectoral composition of workers and micro-entrepreneurs is stable enough that the Intensity measure captures pre-existing exposure, not reform-induced reallocation. To test this, I examine changes in the distribution of workers and firms across eligible CNAE codes before and after the reform and investigate whether there is any evidence of bunching. Figure 7 plots the monthly share of workers employed in MEI-eligible sectors between 2007

and 2015. The series appears relatively stable across the period considered, with fluctuations reflecting short-term variation but no clear upward or downward trend. Few months before the reform, in July 2010, there is a visible short-lived uptick in the share of workers in eligible sectors, followed by a return to pre-reform levels, and if anything, after the reform the share of workers in eligible sectors appears to be lower. This suggests no persistent large-scale reclassification or systematic entry into MEI-eligible sectors in response to the reform, mitigating concerns over endogenous changes in sectoral composition.

Figure 7: Sectoral composition overtime



Note: The figure shows the share of workers employed in MEI-eligible sectors over time, using PME microdata aggregated by month. The vertical red dashed line marks the April 2011 reform. *Source:* Author's calculations based on PME (2007–2015).

6 Results

6.1 Average Effects

Table 2 reports the results for the average impact of the April 2011 MEI reform on formalization rates from the Difference-in-Differences specification (Equation 2 in Section 5.2.1), where the treatment intensity varies by sector and is interacted with an indicator for the post-reform period. The analysis is conducted on an aggregated panel of two-digit CNAE sectors across the six metropolitan regions, observed monthly from 2007 to 2015 (excluding the first quarter of 2011 to account for anticipation effects). Outcomes include the log number of formal micro-entrepreneurs and the share of formal micro-entrepreneurs within the total micro-entrepreneur population.

The estimates indicate a statistically significant sizable positive effect of the reform on both measures of formalization. For the log number of formal micro-entrepreneurs, the interaction term between post-reform and treatment intensity is 0.252 ($p = 0.012$), implying that sectors fully composed of MEI-eligible activities experienced an average increase of approximately 25 percent in the number of formal micro-entrepreneurs relative to non-eligible sectors. For the share of formal micro-entrepreneurs, the estimated coefficient is 0.036 ($p = 0.024$), which translates into a rise of about 3.6 percentage points in fully eligible sectors, or 1.8 points for sectors

Table 2: Average effects of the 2011 reform on formalization (DiD)

	Log(# formal)	**Share formal**
Post × Intensity$_s$	**0.265****	**0.038****
	(0.068)	(0.011)
Observations	22,564	22,564
Clusters: 2-digit CNAE	48	48
Clusters: Metro	6	6
Adj R^2	0.864	0.334
Root MSE	0.479	0.229
Controls (age, sex, race, schooling)	Yes	Yes
Fixed effects	Sector; Metro; Month	Sector; Metro; Month
SEs	Two-way clustered (sector, metro)	Two-way clustered (sector, metro)

Notes: Dependent variables are the log of the number of formal micro-entrepreneurs and the share of formal micro-entrepreneurs at the sector (s) × metro (m) × month (t) level. Controls include sector-level averages of age, age squared, gender, racial composition, and years of education. Fixed effects are at the sector, metro, and month level. Heteroskedasticity-robust standard errors in parentheses are two-way clustered by 2-digit CNAE and metro. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

at median eligibility. These results remain stable when controlling for sector-level demographic composition.

The difference in magnitude between the two outcomes suggests that part of the effect operates through the extensive margin of entry rather than purely through formalization of existing micro-entrepreneurs. An increase of 25 percent in the log number of formal entrepreneurs alongside a modest 3.6 percentage-point rise in the share indicates that the total number of micro-entrepreneurs (the denominator of the share) also expanded post-reform. This pattern is consistent with the creation of new formal businesses under MEI rather than a wholesale transition of existing informal entrepreneurs into formality. This interpretation aligns with prior evidence for Brazil (Rocha et al., 2018) and other Latin American contexts, where simplified regimes often attract new entrants to formal entrepreneurship while only partially reducing informality among incumbents.

These average effects give important insights on the overall impact of the MEI reform on micro-entrepreneurship dynamics and provide a useful benchmark for the analysis of heterogeneous responses presented in the next section.

Figures 8 and 9 visualize the findings by presenting raw trends in formalization outcomes – number and share of formal micro-entrepreneurs respectively – for sectors with high (Intensity > 0.5) versus low (Intensity ≤ 0.5) MEI eligibility exposure, before and after April 2011. Both graphs show parallel pre-trends (formally tested in Section 5.3), while post-reform trajectories diverge between treated and control groups: high-intensity sectors show a sharper rise in the log number of formal businesses (Figure 3) and, despite lower initial formalization shares (reflecting policy targeting), record stronger gains than low-intensity sectors (Figure 4), consistent with reform-driven growth in formality. The figures suggest an uptick in late 2010 - early 2011, consistent with expected anticipatory effects, that in fact are accounted for in the main specification by excluding the first quarter of 2011 from the regression sample.

Figure 8: Raw Trends: Treated vs Control (log formal)

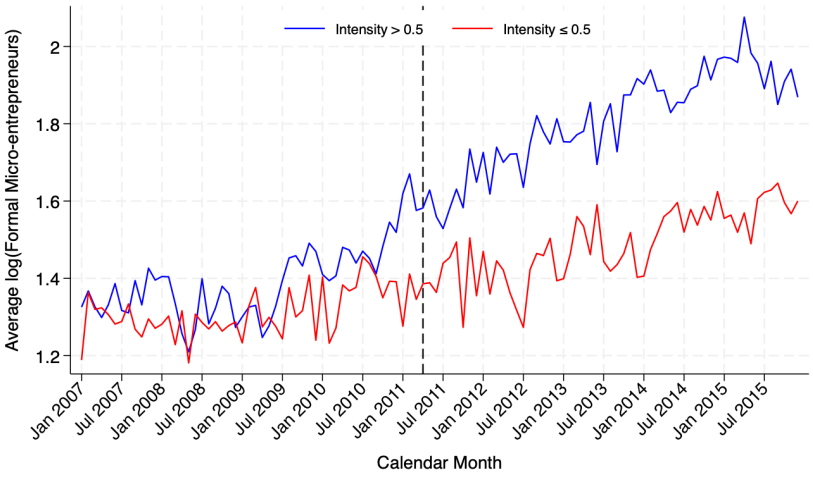
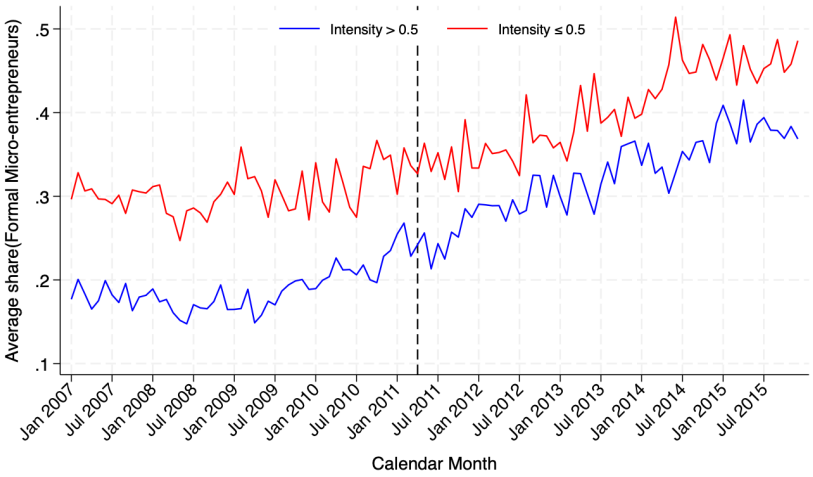


Figure 9: Raw Trends: Treated vs Control (share formal)



Note: Average monthly number of formal micro-entrepreneurs expressed in logarithm (Figure 8) and Average share of formal micro-entrepreneurs out all micro-entrepreneurs (Figure 9) in sectors with high (Intensity > 0.5) versus low (Intensity ≤ 0.5) MEI eligibility exposure. The dashed vertical line denotes the implementation of the MEI tax reduction in April 2011.
Source: Author's calculations based on PME (2007–2015).

6.2 Heterogeneous Effects by Demographic Group: Expected Returns to Social Protection

6.2.1 Maternity Benefits

To examine whether the MEI reform induced larger formalization responses among women of childbearing age – consistent with the incentive effect of access to maternity benefits – I estimate a triple-difference specification interacting post-reform exposure, sectoral treatment intensity, and a demographic indicator. The treated group is defined as women aged 20–40, while the comparison group includes all other micro-entrepreneurs (men and older women). The

estimation is performed on the aggregated panel of sector × metropolitan area × month × demographic group, with outcomes capturing the number and share of formal micro-entrepreneurs (see Equation 3 in Section 5.1.2).

Table 3 reports the results. Contrary to expectations, the interaction term Post × Intensity × Maternity Group is negative and statistically significant for both outcomes. For the number of formal micro-entrepreneurs, the coefficient is -0.168 (p = 0.077), implying that, in fully MEI-eligible sectors, the post-reform increase in formalization was about 16.8 percent lower for women of childbearing age relative to other groups. For the share of formal micro-entrepreneurs, the coefficient is -0.023 (p = 0.037), corresponding to an additional decline of approximately 2.3 percentage points for the treated group in sectors with maximum eligibility. These results hold after controlling for demographic composition and sector-time fixed effects.

Table 3: Differential effects of the 2011 reform (DDD) - Maternity benefits

	Log(# formal)	**Share formal**
Post × Intensity$_s$	0.269***	0.043**
	(0.063)	(0.012)
Post × Intensity$_s$ × Maternity group	**-0.168***	**-0.023****
	(0.076)	(0.008)
Observations	35,176	35,176
Clusters: 2-digit CNAE	48	48
Clusters: Metro	6	6
Adj R^2	0.756	0.289
Root MSE	0.580	0.259
Controls (age, sex, race, schooling)	Yes	Yes
Fixed effects	Sector; Metro; Month	Sector; Metro; Month
SEs	Two-way clustered (sector, metro)	Two-way clustered (sector, metro)

Notes: Estimates from triple-difference models at the sector s × metro m × month t level. Specifications include sector, metro, and month- fixed effects and demographic-composition controls (age, age squared, gender, racial composition, and years of education). Heteroskedasticity-robust standard errors in parentheses are two-way clustered by 2-digit CNAE and metro. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The negative differential effect suggests that, despite the theoretical appeal of maternity benefits as an incentive, the reform did not induce greater formalization among women in childbearing age – and, if anything, their relative responsiveness was lower. This indicates that women of childbearing age may have faced binding constraints – such as lower earnings capacity, higher opportunity costs of compliance, or preference for flexibility – that outweighed the perceived value of maternity benefits. Further discussion of potential mechanisms is provided in Section 6.3.

Figures 10 and 11 plot the gap in formalization rates between women of childbearing age versus men and women outside of childbearing age separately for the high-intensity and low-intensity MEI exposure sectors, with higher values indicating a greater shortfall in formalization among women of childbearing age relative to others. Gaps remain stable prior to the reform and do not narrow in high-intensity sectors afterward, consistent with the negative triple-difference estimates in Table 3.

Figure 10: Gap in number of formal micro-entrepreneurs between women of childbearing age and others, by sector eligibility

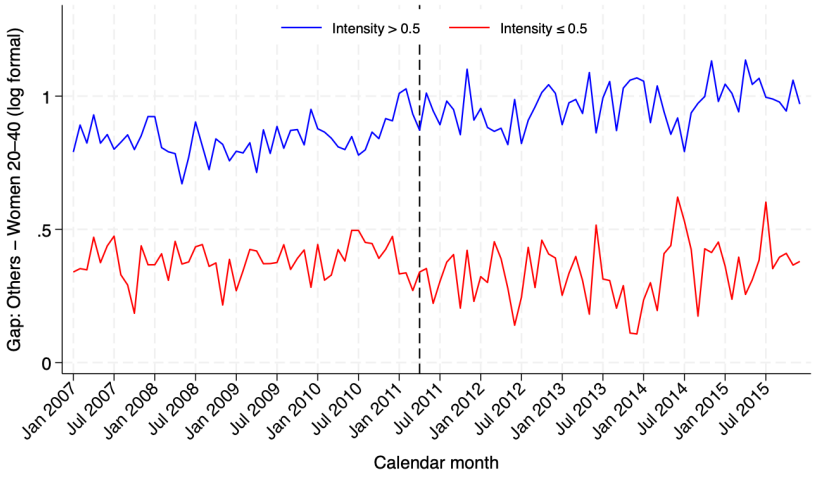
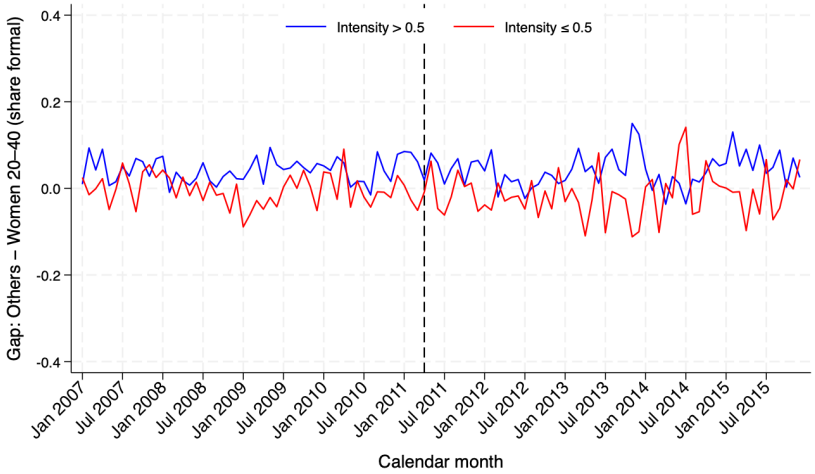


Figure 11: Gap in share of formal micro-entrepreneurs between women of childbearing age and others, by sector eligibility



Note: Difference in the average number of formal micro-entrepreneurs (Figure 10) and difference in the average share of formal micro-entrepreneurs (Figure 9) between two groups within the same exposure category: "Others" (men of all ages plus women aged 40–65) and women aged 20–40. A higher line means women 20–40 are less formalized than Others in that month; a lower line means the gap is smaller. The vertical dashed line marks the April 2011 reform date.
Source: Author's calculations based on PME (2007–2015).

6.2.2 Retirement Benefits

The second potential mechanism of interest is whether older workers nearing retirement responded more strongly to the MEI reform, given the promise of future pension benefits under INSS. To explore this hypothesis, I estimate a triple-difference model similar to that used for the maternity mechanism, defining the treated group as individuals aged 45–55, who are closer to the old-age pension eligibility threshold, and control group asinviduals aged 18-44 and 56-65.

Although theory suggests this category might be more responsive to reduced contribution costs, the results in Table 4 do not support a strong retirement incentive channel. The co-

efficients are small, statistically weak, and inconsistent across outcomes (number of formal micro-entrepreneurs coefficient = -0.070, p = 0.069; share of formal micro-entrepreneurs coefficient = 0.019, p = 0.440). These estimates likely reflect the long time horizon required to access old-age pensions under INSS and limited awareness of benefit accrual among informal workers. As a result, these findings should be interpreted with caution and considered exploratory. They highlight that lowering contribution costs alone is unlikely to influence long-term retirement planning within the time frame of this analysis.

Table 4: Differential effects of the 2011 reform (DDD) - Retirement benefits

	Log(# formal)	**Share formal**
Post × Intensity$_s$	0.269**	0.030*
	(0.069)	(0.012)
Post × Intensity$_s$ × Pension group	**-0.070***	**0.019**
	(0.076)	(0.023)
Observations	39,278	39,278
Clusters: 2-digit CNAE	48	48
Clusters: Metro	6	6
Adj R^2	0.808	0.296
Root MSE	0.4910	0.257
Controls (age, sex, race, schooling)	Yes	Yes
Fixed effects	Sector; Metro; Month	Sector; Metro; Month
SEs	Two-way clustered (sector, metro)	Two-way clustered (sector, metro)

Notes: Estimates from triple-difference models at the sector s × metro m × month t level. Specifications include sector, metro, and month- fixed effects and demographic-composition controls (age, age squared, gender, racial composition, and years of education). Heteroskedasticity-robust standard errors in parentheses are two-way clustered by 2-digit CNAE and metro. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Heterogeneity within the Treated Group: Exploratory Analysis of Maternity Benefits by Individual Characteristics and Potential Channels

To further investigate why the reform did not induce stronger formalization among women of childbearing age, I explore whether the treatment effect varied within this group depending on economic or demographic characteristics . The underlying idea is that even among women of childbearing age, responsiveness to formalization incentives may differ according to constraints and resources at the household or individual level.

These factors include income, working hours, education, number of children under 10 in the household, and household size. Each of these proxies captures a different factor potentially shaping the perceived costs and benefits of formality. For instance, women with higher income or education may be better equipped to internalize the benefits of social protection, whereas those with young children or larger households may face time and caregiving burdens that limit their ability to comply with formalization requirements. Similarly, higher working hours may correlate with more stable employment conditions, which could make the fixed costs of formalization more acceptable.

To assess this, I interact the triple-difference estimator (Post × Intensity × Maternity Group) with a series of pre-determined characteristics measured at the metro × sector × group × month level. To avoid that the policy itself alters characteristics such as income or working hours, creating a risk of post-treatment bias, I use pre-policy baseline values of each heterogeneity

variable. This approach ensures that the heterogeneity dimensions reflect conditions that existed prior to treatment exposure, making them plausibly exogenous with respect to the reform. The regressions are based on the same sample and outcome variables used in Section 6.2.

The results of this exercise are reported in Table 5. First, income emerges as a robust predictor of differential responses, both in the log specification and in the share specification, with the interaction term between Post × Intensity × Maternity group × Income being positive and statistically significant. This indicates that higher-income women within the maternity group were more likely to formalize after the reform relative to lower-income counterparts, suggesting an "affordability channel": economic resources play a critical role in enabling women to bear the costs of registration and in increasing the salience of social security benefits.

Second, household size appears to matter in the opposite direction. The estimates using share formal as the outcome show a negative and statistically significant coefficient for the interaction with household size, suggesting that women from larger households were less likely to formalize in response to the reform. This pattern is consistent with the idea that care and dependency burdens in larger households may increase opportunity costs and liquidity constraints, reducing the appeal of formalization even in the presence of social protection benefits.

By contrast, no systematic heterogeneity is detected along the dimensions of education, working hours, or the number of young children in the household. These coefficients remain statistically insignificant across specifications, suggesting that these characteristics do not meaningfully condition responsiveness to the reform within the scope of this policy and dataset.

Table 5: Heterogeneous Treatment Effects by Maternity Group Characteristics

	(1) Log(Formal)	(2) Share(Formal)
Post × Intensity$_s$ × Maternity × Educ. Years	0.030	-0.000
	(0.050)	(0.013)
Post × Intensity$_s$ × Maternity × ln(Income)	**0.221****	**0.080*****
	(0.067)	(0.016)
Post × Intensity$_s$ × Maternity × Hours Usual	0.019	0.009*
	(0.011)	(0.004)
Post × Intensity$_s$ × Maternity × N. Child <10	0.046	0.075
	(0.490)	(0.071)
Post × Intensity$_s$ × Maternity × Family Size	-1.979	**-0.901****
	(1.653)	(0.312)
Observations	35,129	35,129

Notes. The table reports coefficient estimates for the quadruple interaction term Post×Intensity×Maternity Group×Z, where Z denotes the heterogeneity variable indicated in each row. Pre-policy baseline values are used, to ensure the heterogeneity dimension reflects structural differences in baseline economic conditions rather than treatment itself. The dependent variables are the log number of formal micro-entrepreneurs (Column 1) and the share of formal micro-entrepreneurs (Column 2) at the sector × metropolitan area × month × demographic group level. All regressions include controls for age, age squared, sex, race, and education (except when education is the heterogeneity variable), as well as fixed effects for sector, metropolitan area, and month. Standard errors, clustered at the sector × metropolitan area level, are reported in parentheses. *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In sum, the analysis highlights that maternity benefits under the MEI reform did not operate uniformly across all women of childbearing age. Instead, their effectiveness appears conditioned by economic capacity and family responsibilities, underscoring the importance of household context in shaping the returns to social protection policies.

6.4 Robustness Checks

In this section I conduct multiple robustness checks to assess the sensitivity of the main results to alternative model specifications and sample definitions. All empirical tests are reported in section C of the Appendix.

First, I assess sensitivity to the treatment intensity variable ($Intensity_s$). Instead of relying on the continuous treatment intensity measure, I construct two binary indicators: one that equals one if a sector's intensity is above the median ($Intensity_s > 0.5$), and one that equals one if a sector is above the 90th percentile ($Intensity_s > 0.90$). The results, reported in Tables A.4 and A.5, show that the binary eligibility thresholds yield estimates that are qualitatively similar to those obtained with the continuous measure. In both DiD and DDD specifications, the post-treatment effect remains positive and statistically significant, and the heterogeneous effect for the maternity group is negative, consistent with the baseline specification. This indicates that the results are not sensitive to the choice of functional form, reinforcing the robustness of the results.

Moreover, I assess the sensitivity of the results to anticipatory effects, by relaxing the initial sample restriction and including the first quarter of 2011, previously excluded. Across both specifications, the estimated coefficients remain positive, statistically significant, and similar in magnitude, indicating that the main results are not driven by the exclusion of early-2011 observations (Tables A.6 and A.7).

To account for possible omitted variables, I re-estimate the models using stricter sets of fixed effects. First, I include metropolitan × month fixed effects to absorb any metro-level shocks that vary month to month (e.g., local business cycles, enforcement changes). Second, I introduce sector × metropolitan fixed effects to control for time-invariant sorting of sectors across cities. In both cases, the treatment coefficients remain robust, with effect sizes and significance levels closely aligned with the baseline specification (Tables A.8 and A.9).

Finally, I test robustness to different standard error specifications. While my main regressions cluster at the metro × sector level, I re-estimate models with clustering at the metropolitan area level and at the industry level separately (Table A.10 and A.11). The results are virtually unchanged in both magnitude and significance, confirming that the findings are not sensitive to the choice of clustering scheme.

7 Limitations

Several limitations should be acknowledged. A first concern is the potential for strategic reclassification of employment relationships—commonly referred to in Brazil as *pejotização* (OECD, 2020)—whereby firms encourage or oblige workers to register as MEIs to reduce labor costs and obligations. Such practices could artificially inflate formalization rates without reflecting genuine transitions from informality to autonomous entrepreneurship.

Second, while the analysis captures transitions into formality, it does not address their durability. The research studies up to 4 years after the policy, with no estimation of longer-run effects. Short-term incentives, such as eligibility for maternity benefits, may induce temporary registrations that do not translate into sustained behavioral change, raising the question of

whether observed effects represent lasting integration into the formal system or short-lived spikes in registration.

There are also measurement constraints stemming from the data. PME does not include a direct identifier for MEI participation, so formalization is inferred from reported employment status and contributions to INSS, which introduces potential measurement error. Moreover, PME provides sectoral information only at the two-digit CNAE level, while eligibility is defined at a more granular level. To bridge this gap, I rely on a treatment intensity proxy constructed with PNAD data, which may attenuate the estimated effects due to sectoral misclassification. In addition, due to limitations in the dataset, PME's rotating panel design does not allow for the tracking of individuals across time, preventing direct observation of dynamic transitions from informal to formal status and restricting the analysis to aggregated outcomes.

Interpretation of the retirement mechanism also warrants caution. While the analysis assumes that mid-aged workers (45–55) may respond to reduced contribution costs due to proximity to pension eligibility, INSS retirement benefits require at least 15 years of contributions. The expected returns are therefore highly uncertain within the time horizon of this study, making it difficult to disentangle genuine pension-driven incentives from other age-related factors.

Furthermore, it is important to note that the analysis focuses on Brazil's six largest metropolitan areas—São Paulo, Rio de Janeiro, Belo Horizonte, Porto Alegre, Recife, and Salvador, which together represent more than 40% of the population (IBGE, 2025). While this urban focus ensures robust statistical power and captures the settings where administrative enforcement and outreach were strongest and where self-employment is most concentrated, nevertheless it limits the generalizability of the results to rural or less formalized contexts, where informality dynamics may differ and enforcement capacity is weaker.

Finally, the analysis is restricted to micro-entrepreneurs, the direct target of the MEI programme, and does not consider potential spillover effects on other categories of workers.

8 Conclusions

This study has examined the role of contributory social security as a driver of labor market formalization in Brazil, focusing on the 2011 reform of the MEI programme that halved the monthly social security contribution required from registered micro-entrepreneurs. By exploiting this policy change as a quasi-natural experiment, I have been able to identify the causal impact of lowering contribution costs on formalization outcomes, both on average and across demographic groups differentially likely to value specific social protection benefits.

The results demonstrates that reducing the cost of social security contributions led to a significant increase in the number and share of formal micro-entrepreneurs on average, in line with theories emphasizing the importance of cost–benefit calculations in the formalization decision. Yet, the evidence also suggests that much of the growth occurred at the extensive margin—through the creation of new formal businesses—rather than widespread transitions from informality among incumbents.

Contrary to the initial hypothesis, the heterogeneity analysis reveals that women of childbearing age did not respond more positively to the reform despite their eligibility for paid maternity leave. In fact, their relative formalization rates declined compared to men and older

women. Exploratory analyses suggest that income constraints and household responsibilities acted as significant barriers, preventing poorer women from translating eligibility into uptake. The analysis of mid-aged workers provided no consistent evidence of a retirement channel, which is consistent with the long contribution horizon required to access pensions and the relatively low salience of such benefits for workers facing income insecurity and short-term planning horizons.

Taken together, these findings suggest that income is the decisive factor mediating whether social protection functions as an incentive to formalize. Contribution reductions can encourage entry into formality, but unless they are calibrated to the capacity of low-income entrepreneurs to pay, large segments of the informal workforce will remain beyond reach. This underscores the importance of designing differentiated contribution schedules or progressive schemes that scale costs to income, ensuring that the groups most in need are not systematically excluded.

Together, these findings provide new evidence on the complex interaction between social protection and informality. They show that contributory schemes can act as incentives to formalize, but only when workers are both able and willing to bear the cost of participation. The results also suggest that the design of formalization policies must go beyond lowering contributions to address structural barriers that disproportionately affect vulnerable groups, particularly women —such as caregiving burdens and limited access to credit.

In sum, this study shows that contributory social protection can act as a pathway to formalization, but only if affordability constraints are addressed. Without doing so, reforms risk deepening divides between those able to participate and those who remain excluded from both formality and its protections.

9 References

Aterido, R., Hallward-Driemeier, M. C., & Pages, C. (2011). Does expanding health insurance beyond formal-sector workers encourage informality? Measuring the impact of Mexico's Seguro Popular (Policy Research Working Paper No. WPS 5785). World Bank.

Azuara, O., & Marinescu, I. (2013). Informality and the expansion of social protection programmes: Evidence from Mexico. *Journal of Health Economics*, 32(5), 938-950.

Barrientos, A. (2011). Social protection and poverty. *International Journal of Social Welfare*, 20(3), 240–249.

Bérgolo, M., & Cruces, G. (2011). Labor informality and the incentive effects of social security: Evidence from a health reform in Uruguay (Working Paper). Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS).

Besley, T., & Persson, T. (2013). Taxation and development. In A. J. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics* (Vol. 5, pp. 51–110). Elsevier.

Bosch, M., & Campos-Vazquez, R. M. (2014). The trade-offs of welfare policies in labor markets with informal jobs: The case of the "Seguro Popular" programme in Mexico. *American Economic Journal: Economic Policy*, 6(4), 71-99.

Bruhn, M. (2011). Entry regulation and the formalization of microenterprises in developing countries. *Labour Economics*, 18(Suppl. 1), S12-S25.

Camacho, A., Conover, E., & Hoyos, A. (2014). Effects of Colombia's social protection system on workers' choice between formal and informal employment. *World Bank Economic Review*, 28(3), 446-466.

Canelas, C., & Niño-Zarazúa, M. (2022). Social protection and the informal economy: What do we know? *Journal of Development Studies*, 58(4), 569-588.

De Giorgi, G., & Rahman, A. (2013). SME registration evidence from a randomized controlled trial in Bangladesh. *Economics Letters*, 120(3), 573-578.

Fajnzylber, P., Maloney, W. F., & Montes-Rojas, G. V. (2011). Out of the shadows? Revisiting the impact of the Brazilian SIMPLES programme on firms' formalization rates. *Journal of Development Economics*, 94(2), 284-295.

Firpo, S. Portella, A., (2021). Informal and small: How labor market institutions affect inequality, income volatility and labor productivity in Brazil. *UNDP LAC Working Paper No.22.*

Gerard, F., & Naritomi, J. (2021). Job displacement insurance and (the lack of) consumption-smoothing. *American Economic Review*, 111(3), 899-942.

Gruber, J. (1994). The incidence of mandated maternity benefits. *American Economic Review*, 84(3), 622-641.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4), 1091-1119.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), 1403-1448.

International Labour Organization. (2015). *Recommendation No. 204: Concerning the transition from the informal to the formal economy, adopted by the Conference at its one hundred and fourth session, Geneva, 12 June 2015.* International Labour Organization.

International Labour Organization (ILO). (2017). *World employment social outlook: Trends* 2017.

International Labour Organization (ILO). (2018). *Women and men in the informal economy: A statistical picture* (3rd ed.). International Labour Office.

International Labour Organization (ILO). (2023). *Employment and informality in Latin America and the Caribbean.*

International Labour Organization (ILO). (2024). *Innovative approaches to promoting the transition to formality in Latin America and the Caribbean.* International Labour Office.

Kaplan, D. S., Piedra, E., & Seira, E. (2011). Entry regulation and business start-ups: Evidence from Mexico. *Journal of Public Economics*, 95(11-12), 1501-1515.

Kugler, A., & Kugler, M. (2009). Labor market effects of payroll taxes in developing countries: Evidence from Colombia. *Economic Development and Cultural Change*, 57(2), 335-358.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109-126.

Levy, S. (2008). *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico.* Brookings Institution Press.

Liepmann, H., & Pignatti, N. (2021). Welfare effects of unemployment benefits when informality is high. *Labour Economics,* 70, 101970.

Maloney, W. F. (2004). Informality revisited. *World Development,* 32(7), 1159-1178.

Marchese, M. (2021). Preferential tax regimes for MSMEs: Operational aspects, impact evidence and policy implications (OECD Working Papers on SMEs and Entrepreneurship No. 22). OECD Publishing.

Monteiro, J. C. M., & Assunção, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. J*ournal of Development Economics*, 99(1), 105-115.

OECD. (2020). *SME and entrepreneurship policy in Brazil 2020.* OECD Publishing.

Olden, A., & Møen, J. (2022). Triple difference estimator. *Journal of Econometrics*, 230(1), 51-69.

Pereda, P., et al. (2022). Formalization of female micro-entrepreneurs in Brazil (Working Paper). *Inter-American Development Bank*.

Rocha, R., Ulyssea, G., & Rachter, L. (2018). Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, 134, 28-49.

Schatan, R., Benítez, J. C., Coelho, I., Madariaga, J. (2022). Peru: Tax regime for small taxpayers and special economic zones. *IMF Technical Assistance Report*, International Monetary Fund.

Torm, N., & Oehme, J. (2024). Social protection and formalization in low- and middle-income countries: A scoping review of the literature. *World Development*, 180, 106376.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8), 2015-2047.

Appendices

A. Treatment Intensity Variable

Detailed Explanation of Variable Construction: The MEI policy defines eligibility at the 7-digit CNAE subclass level. At this level, each subclass corresponds to a very specific occupation or activity (e.g., *"retail bakery without production"*, *"taxi driver"*, *"beauty salon"*). Some subclasses are explicitly listed in the government's administrative regulations as MEI-eligible, while others are not. However, the PME dataset only provides sectoral identifiers at the 2-digit CNAE division level (e.g., *"Food products manufacturing"*, *"Retail trade"*, *"Education"*). To bridge this mismatch, I construct an aggregated measure of sectoral exposure to the policy by calculating, for each 2-digit sector, the share of underlying 7-digit subclasses that are MEI-eligible.

Formally, for each 2-digit sector s, the treatment intensity is defined as:

$$Intensity_s = \frac{\text{Number of MEI-eligible subclasses within sector } s}{\text{Total number of subclasses within sector } s}.$$

This gives a continuous variable between 0 and 1, where:

- $Intensity_s = 0$ means no activity in that sector is eligible for MEI,
- $Intensity_s = 1$ means all activities in that sector are eligible,
- Values in between reflect partial exposure, depending on how many detailed occupations fall under MEI.

Suppose the 2-digit sector "Manufacture of Food Products" (CNAE 15) contains 100 distinct 7-digit subclasses in the CNAE classification. Out of these, 99 subclasses are listed as eligible under the MEI regulation. Then:

$$Intensity_{\text{Food}} = \frac{99}{100} = 0.99.$$

This high intensity (close to 1) reflects that almost all food-related microentrepreneurs could formalize under MEI. This sector-level measure captures the *degree of exposure* to the reform, not just a binary eligible/ineligible assignment. It allows me to exploit cross-sector variation in intensity to identify heterogeneous treatment effects in a difference-in-differences framework. Because intensity is continuous, results are not sensitive to an arbitrary cutoff. In robustness checks, I also replace it with binary thresholds (e.g., above/below median, above 90th percentile) and find similar patterns.

Table A.1 reports the calculated intensity values for each 2-digit CNAE sector included in the analysis, alongside illustrative examples of MEI-eligible sub-industries. This table clarifies the mapping from detailed eligibility rules — defined at the 7-digit CNAE level in administrative regulations—to the aggregated sector classifications observed in PME. By showing both the numeric intensity and representative eligible activities, it helps illustrate the heterogeneity in exposure leveraged for identification. The last column indicates whether the sector was included in the analysis - due to labour market -comparability or not.

Table A.1: APPENDIX TABLE A.1: Mapping of CNAE Sectors to Treatment Intensity and Eligible Sub-Industries

CNAE 2-digit	Intensity	Sector description	Included
00	0.0000000	Unspecified activities	no
01	0.0843027	Agriculture, livestock and related services	no
02	0.0000000	Forestry and logging, related services	no
05	0.0198347	Fishing and aquaculture, related services	no
10	0.0000000	Coal mining	no
11	0.0000000	Oil and gas extraction and related services	no
12	0.0000000	Radioactive mineral extraction	no
13	0.0000000	Metal ore mining	no
14	0.0000000	Non-metallic mineral mining	no
15	0.9942396	Manufacture of food products and beverages	yes
16	1.0000000	Manufacture of tobacco products	yes
17	1.0000000	Manufacture of textiles	yes
18	1.0000000	Manufacture of wearing apparel and accessories	yes
19	1.0000000	Leather preparation and manufacture of leather goods, travel goods and footwear	yes
20	1.0000000	Manufacture of wood products	yes
21	1.0000000	Manufacture of pulp, paper and paper products	yes
22	1.0000000	Publishing, printing and reproduction of recordings	yes
23	0.0000000	Coke ovens	yes
24	0.2413793	Manufacture of chemicals	yes
25	0.0000000	Manufacture of rubber and plastics	yes
26	0.7971014	Manufacture of non-metallic mineral products	yes
27	0.0000000	Basic metals	yes
28	0.2138728	Fabricated metal products (except machinery and equipment)	yes
29	0.0000000	Machinery and equipment	yes
30	0.0000000	Computers and electronic equipment	yes
31	0.0000000	Electrical machinery and apparatus	yes
32	0.0000000	Electronic and communication equipment	yes
33	0.0000000	Medical, precision and optical instruments; industrial automation; clocks and watches	yes
34	0.0000000	Manufacture and assembly of motor vehicles, trailers and bodies	yes
35	0.0000000	Other transport equipment	yes
36	1.0000000	Furniture and other industries	yes
37	0.0000000	Recycling	yes
40	0.0000000	Production and distribution of electricity, gas and water	yes
41	0.0000000	Collection, treatment and distribution of water	yes
45	1.0000000	Construction	yes
50	0.8694981	Trade and repair of motor vehicles and motorcycles; retail of fuels	yes
53	0.3671159	Trade intermediaries; trade and repair of personal and household goods	yes
55	1.0000000	Hotels and restaurants	yes
60	0.4461667	Land transport	yes
61	0.0000000	Water transport	yes
62	0.0000000	Air transport	yes
63	0.5958549	Auxiliary transport activities and travel agencies	yes
64	1.0000000	Post and telecommunications	yes
65	0.0000000	Financial intermediation (excluding insurance and pensions)	yes
66	0.0000000	Insurance and pension funding	yes
67	0.0000000	Activities auxiliary to financial intermediation	yes
70	0.0000000	Real estate activities	yes
71	0.8782609	Renting of vehicles, machinery, equipment and personal goods	yes
72	0.3981265	Computer and related activities	yes
73	0.0000000	Research and development	yes
74	0.3182927	Business services, mainly to firms	yes
75	0.0000000	Public administration, defense and social security	no
80	0.8975610	Education	yes
85	0.0204082	Health and social services	yes
90	0.0000000	Urban cleaning and sewerage, related activities	yes
91	0.0000000	Associative activities	yes
92	0.9843972	Recreational, cultural and sporting activities	yes
93	0.8825895	Personal services	yes
95	0.0000000	Domestic services	no
99	0.0000000	International organizations and extraterritorial institutions	no

B. Identification Assumption Checks

Below are the estimates for the falsification tests I conduct using pre-reform periods only (2007–2010), assigning pseudo-reform dates to January 2008, January 2009, and January 2010. I find no significant coefficients for formalization outcomes, indicating that the estimated effects in the main analysis are not driven by pre-existing differential trends across treatment and control groups. This supports the validity of the identification strategy and reinforces the assumption that sectoral exposure to MEI eligibility is exogenous to underlying formality dynamics.

Table A.2: Placebo Tests: Effect of MEI Reform in Pre-Treatment Periods - Average effects

	January 2008		January 2009		January 2010	
	(1) Log	(2) Share	(3) Log	(4) Share	(5) Log	(6) Share
Post_placebo × Intensity	0.016	-0.006	0.047	-0.009	0.058	0.003
	(0.039)	(0.014)	(0.027)	(0.017)	(0.038)	(0.021)
Observations	10,261	10,261	10,261	10,261	10,261	10,261
R-squared	0.851	0.342	0.850	0.342	0.850	0.342

Notes. Robust standard errors in parentheses. The table reports falsification tests using only pre-reform periods (2007–2010). Coulumns (1), (3), (5) report estimates for the number of formal micro-entrepreneurs (expressed in logarithm), while columns (2), (4), (6) present results for the share of formal micro-entrepreneurs.

Table A.3: Placebo Tests: Heterogeneous Effects for Maternity Group (Pre-Treatment Periods)

	January 2008		January 2009		January 2010	
	(1) Log	(2) Share	(3) Log	(4) Share	(5) Log	(6) Share
Post_placebo × Intensity × Maternity group	0.004	0.003	-0.012	-0.013	0.005	0.007
	(0.074)	(0.032)	(0.080)	(0.032)	(0.047)	(0.026)
Observations	16,098	16,098	16,098	16,098	16,098	16,098
R-squared	0.739	0.295	0.739	0.295	0.739	0.295

Notes. Robust standard errors in parentheses. The table reports falsification tests for heterogeneous treatment effects by maternity group using only pre-reform periods (2007–2010). Coulumns (1), (3), (5) report estimates for the number of formal micro-entrepreneurs (expressed in logarithm), while columns (2), (4), (6) present results for the share of formal micro-entrepreneurs.

C. Robustness Checks

C.1 Alternative Treatment Eligibility – Binary Thresholds

Table A.4: Robustness: Binary Eligibility Thresholds in DiD (Median and 90th Percentile)

	Intensity > 0.5		**Intensity** > 0.9	
	(1) Log(Formal)	(2) Share(Formal)	(3) Log(Formal)	(4) Share(Formal)
Post × Eligible	**0.205****	**0.032****	**0.148***	**0.019***
	(0.058)	(0.009)	(0.064)	(0.009)
Observations	22,564	22,564	22,564	22,564
R-squared	0.865	0.339	0.864	0.338

Notes. DiD estimates using binary eligibility indicators: *Eligible* = 1 if sector intensity exceeds the median (cols. 1–2) or the 90th percentile (cols. 3–4). Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls and fixed effects as in the baseline. SEs clustered at the CNAE 2-digit × metro level.

Table A.5: Robustness: Binary Eligibility in DDD (Maternity Group)(Median and 90th Percentile

	Intensity > 0.5		**Intensity** > 0.9	
	(1) Log(Formal)	(2) Share(Formal)	(3) Log(Formal)	(4) Share(Formal)
Post × Eligible	0.209**	0.034**	0.147*	0.022**
	(0.054)	(0.010)	(0.062)	(0.009)
Post × Eligible × Maternity group	**-0.141**	**-0.015**	**-0.155**	**-0.026****
	(0.074)	(0.009)	(0.081)	(0.009)
Observations	35,176	35,176	35,176	35,176
R-squared	0.752	0.292	0.753	0.292

Notes. DDD estimates with binary eligibility (median and 90th percentile). The interaction with the maternity group captures heterogeneous effects for women of childbearing age. Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls and fixed effects as in the main DDD; SEs clustered at CNAE 2-digit × metro.

C.2 Robustness to Anticipatory Effects

Table A.6: Robustness to Anticipatory Effects: Including 2011Q1 (DiD)

	(1) Log(Formal)	(2) Share(Formal)
Post × Intensity	**0.252****	**0.036****
	(0.065)	(0.011)
Observations	23,201	23,201
R-squared	0.865	0.338

Notes. Same specification as the baseline DiD but relaxing the exclusion of early 2011 and including 2011Q1. Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls and fixed effects as in the baseline; clustering at CNAE 2-digit × metro.

Table A.7: Robustness to Anticipatory Effects: Including 2011Q1 (DDD - Maternity Group)

	(1) Log(Formal)	(2) Share(Formal)
Post × Eligible	0.256***	0.040**
	(0.060)	(0.011)
Post × Eligible × Maternity group	**-0.156***	**-0.019****
	(0.074)	(0.006)
Observations	36,183	36,183
R-squared	0.757	0.292

Notes. DDD with binary eligibility, adding 2011Q1 back to test for anticipatory effects. Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls include sector-level averages of age, age squared, gender, racial composition, and years of education; fixed effects for sector, metro, and month. SEs clustered at CNAE 2-digit × metro.

C.3 Alternative Fixed Effect Specifications

Table A.8: DiD Results Under Alternative Fixed Effects

	Log(Formal)			Share(Formal)		
	(1) baseline	(2) FE–A	(3) FE–B	(4) baseline	(5) FE–A	(6) FE–B
Post × Intensity	**0.265****	**0.246****	**0.248*****	**0.038****	**0.034****	**0.036****
	(0.068)	(0.066)	(0.060)	(0.011)	(0.012)	(0.011)
Observations	23,201	23,201	23,201	23,201	23,201	23,201
R-squared	0.864	0.870	0.900	0.335	0.358	0.366

Notes. Each column reports a separate DiD regression of the stated outcome on $Post \times Intensity_s$ with controls (age, age^2, male, Black/Brown, education years). FE–A includes *sector* FE and *metro*×*month* FE; FE–B includes s*ector*×*metro* FE and *month* FE. Robust standard errors clustered at CNAE 2-digit and metro shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: DDD Results (Heterogeneity by Maternity Group) Under Alternative Fixed Effects

	Log(Formal)			**Share(Formal)**		
	(1) baseline	(2) FE–A	(3) FE–B	(4) baseline	(5) FE–A	(6) FE–B
Post × Intensity	0.269	0.252	0.253	0.043	0.039	0.040
	(0.063)	(0.061)	(0.056)	(0.012)	(0.012)	(0.010)
Post × Intensity × Maternity	**-0.168***	**-0.159***	**-0.167***	**-0.023****	**-0.019****	**-0.018****
	(0.076)	(0.074)	(0.074)	(0.008)	(0.007)	(0.005)
Observations	36,183	36,183	36,183	36,183	36,183	36,183
R-squared	0.756	0.762	0.784	0.289	0.306	0.315

Notes. Each column reports a DDD regression interacting *Post*, $Intensity_s$, and an indicator for the maternity group (women of childbearing age). Controls: age, age^2, male, Black/Brown, education years. FE–A includes *sector* FE and *metro*×*month* FE; FE–B includes *sector*×*metro FE* and *month* FE. Robust standard errors clustered at CNAE 2-digit and metro shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Alternative Clustering Specifications

Table A.10: DiD Results Under Alternative SE Clustering

	Log(Formal)			Share(Formal)		
	(1) Baseline	(2) Sector	(3) Metro	(4) Baseline	(5) Sector	(6) Metro
Post × Intensity	**0.265****	**0.252*****	**0.252*****	**0.038****	**0.036*****	**0.036****
	(0.068)	(0.057)	(0.045)	(0.011)	(0.012)	(0.011)
Observations	23,201	23,201	23,201	23,201	23,201	23,201
R-squared	0.864	0.864	0.864	0.335	0.337	0.337

Notes. Each column reports a separate DiD regression of the stated outcome on $Post \times Intensity_s$ with controls (age, age^2, male, Black/Brown, education years). Baseline robust standard errors are clustered at CNAE 2-digit and metro. Robustness checks test for standard errors clustered solely at 2-digit sector level (columns (2),(5)), and for standard errors clustered only at metropolitan region level (columns (3) and (6)) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: DDD Results (Maternity Group) Under Alternative SE Clustering

	Log(Formal)			**Share(Formal)**		
	(1) Baseline	(2) Sector	(3) Metro	(4) Baseline	(5) Sector	(6) Metro
Post × Intensity	0.269	0.252	0.253	0.043	0.039	0.040
	(0.063)	(0.061)	(0.056)	(0.012)	(0.012)	(0.010)
Post × Intensity × Maternity	**-0.168***	**-0.156****	**-0.156****	**-0.023****	-0.019	**-0.019***
	(0.076)	(0.070)	(0.051)	(0.008)	(0.007)	(0.009)
Observations	36,183	36,183	36,183	36,183	36,183	36,183
R-squared	0.756	0.757	0.757	0.289	0.292	0.292

Notes. Each column reports a DDD regression interacting *Post*, $Intensity_s$, and an indicator for the maternity group (women of childbearing age). Controls: age, age^2, male, Black/Brown, education years. Baseline FE include textitsector, textitmetropolitan area, textitdate separately; FE–A includes *sector* FE and *metro*×*month* FE; FE–B includes *sector*×*metro FE* and *month* FE. Robust standard errors clustered at CNAE 2-digit and metro shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.