Constructing Structural Change*

Julieta Caunedo   Mayara Felix   Kristina Manusheva

August 6, 2026
Preliminary and in progress

Abstract

This paper studies how non-agricultural labor demand shapes the organization of agricultural production. Using municipality-level panel data from Brazil, we combine two complementary quasi-experimental designs—staggered entry of large non-agricultural firms and variation in a federal housing program (Minha Casa Minha Vida)—to show that construction-sector temporary labor demand raises agricultural wages, reduces agricultural employment, and triggers mechanization, farm consolidation, and productivity growth. Construction entries generate these effects; manufacturing entries of comparable size do not. The difference stems from sectorial skill composition: agriculture and construction share a predominantly low-skill labor force, while modern Brazilian manufacturing draws from a higher-skill pool. A general-equilibrium model featuring endogenous mechanization with a fixed cost of technology adoption that induces irreversibility rationalizes the patterns. Once farms invest in mechanizing a set of production tasks, such investment persists as factor factor prices revert to their initial levels. Since the impact on agriculture depends on the induced shift in wages as well as the suitability of mechanized practices, two dimensions of heterogeneity further support our mechanisms: informality attenuates the agricultural response in line with a Lewis-type labor buffer, while the mechanization response is largest at intermediate terrain steepness. We find evidence of farm consolidation even where mechanization response is weaker.

Keywords: Structural Transformation, Agricultural Organization, Mechanization, Labor Demand.

JEL Codes: O13; O14; O47; Q12; R23.

---

*Caunedo: Cornell University and CEPR, email: julieta.caunedo@cornell.edu; Felix: Yale University and NBER, email: mayara.felix@yale.edu; Manysheva: Columbia University, email: km3924@columbia.edu. We thank Andrew Bernard, Kevin Donovan, Michael Peters, Markus Poschke, Mark Rosenzweig, and audiences at Columbia University, Princeton University, Barcelona Summer Forum, SED Winter ’24, LACEA ’25, IGC PEDL ’25, SED ’26, the FLD Workshop ’26, the Richmond Fed, and the Minneapolis Fed for their comments and suggestions. Guilherme Bayma, Madhav Malhotra, Sofía Pontigo and Ekaterina Zubova provided excellent research assistance. This project was supported by the CAF Development Grant of Latin America and the Caribbean.

1 Introduction

A central fact of economic development is the reallocation of labor out of agriculture. Classical frameworks attribute this process to productivity gains in agriculture that release workers (Schultz, 1953; Gollin et al., 2007), or to rising productivity and wages in non-agriculture that attract them (Harris and Todaro, 1970). In practice, both forces operate simultaneously and their relative importance varies over time and across countries (Alvarez-Cuadrado and Poschke, 2011). This paper studies the mechanisms and heterogeneity underlying the process of labor reallocation and agricultural reorganization —mechanization, land consolidation—. Are all positive labor-demand shocks in non-agriculture effective in spurring such reorganization? If not, why? And what features of the local economy, prior to the shock, determine the nature of the response?

We study these questions empirically and quantitatively in the context of 21st-century structural transformation in Brazil. A positive non-agricultural labor-demand shock does reorganize agriculture: it raises agricultural wages, draws workers out of farming, and induces mechanization, farm consolidation, and productivity growth among the farms that survive. But not every shock does so. The skill composition of the labor-demand shock is the key determinant of these effects—construction-driven demand, which competes for the same low-education workers as agriculture, reorganizes farming, whereas manufacturing-driven labor-demand shocks of comparable size do not. The response is further shaped by features of the local economy: it is muted where a large informal sector buffers the labor market, and where terrain limits the scope for mechanization. Despite the shift in labor demand being temporary, the reorganization of production is persistent. A general-equilibrium model with irreversible mechanization rationalizes these patterns and connects the farm-level evidence to aggregate structural change.

We use rich municipality-level panel data from Brazil covering agricultural censuses (1997, 2006, 2017), matched employer-employee records (RAIS), and population censuses (1991, 2000, 2010). Brazil offers an ideal setting for these questions: there is substantial cross-municipality heterogeneity in both non-agricultural employment growth and agricultural structure, and rich administrative records that allow us to track formal-sector employment and wages at annual frequency alongside decadal agricultural reorganization.

We use two complementary quasi-experimental designs. The first exploits the staggered entry of large non-agricultural firms into local labor markets between 1998 and 2017. Using the De Chaisemartin and d'Haultfoeuille (2020) staggered difference-in-differences estimator, we compare municipalities that experienced a large entry "shock"—defined as a new establishment employing at least 200 workers and representing more than 0.1 percent of its microregion's population at entry—with municipalities in microregions that never received such entries. The annual frequency of RAIS lets us estimate event-study paths covering five pre-entry and ten post-entry years, so we can verify parallel pre-trends and trace wage and employment dynamics over a full decade. The second design exploits Brazil's Minha Casa Minha Vida

(MCMV) program—the largest federal low-income housing initiative in the country's history, which contracted roughly 3.8 million housing units between 2009 and 2014 (about 1.9 million in the most-subsidized *Faixa 1* tier). Because the spatial allocation of MCMV contracts was partly governed by political-economy factors—in particular, mayoral alignment with the federal governing coalition—we follow Simoni Junior and Dias (2025) and use these political predictors to isolate the component of local construction activity driven by program allocation rather than by pre-existing conditions. Since the agricultural census is observed only at decade intervals, the MCMV design is well suited to detecting decadal agricultural reorganization, while the large-entry design supplies the high-frequency dynamics that anchor the interpretation of the mechanism.

The large-entry results reveal two facts that are central to interpreting our evidence. First, a large non-agricultural entry tightens the local labor market: formal non-agricultural employment at entry epicenters rises by about 10 log points in the year of entry relative to never-entered municipalities, with parallel pre-trends, and the real wages of less-educated workers rise and continue to grow over the following decade. Second, these effects are driven entirely by construction and real-estate entries, which raise non-agricultural employment by about 13 log points at entry and the real wage of primary-education workers by 12 log points, rising further to 52 log points above never-entered controls a decade later; manufacturing entries of comparable size generate employment and wage coefficients near zero and indistinguishable from pre-entry trends.

A defining feature of these construction-driven shocks is their persistence: the real wage of primary-education workers at construction epicenters does not revert, but stands 52 log points above never-entered municipalities a decade after entry, still growing rather than fading. The employment response, by contrast, is more concentrated around entry and less precisely estimated at long horizons.

Where predicted construction intensity is higher, agricultural employment fell more, agricultural wages rose more, farms adopted more tractors, farms consolidated into larger operations, and agricultural output per worker grew faster over 2006–2017, with no analogous gradient in the pre-program period 1997–2006. All five outcomes move in the same direction simultaneously, consistent with a single mechanism: a rise in the relative cost of agricultural labor makes labor-intensive farming less profitable, inducing substitution of capital for labor, exit of marginal farms, consolidation of land in surviving operations, and productivity gains from farm selection and capital intensification.

What drives the contrast across sectors is not the size of the entry but the skill of the workers it draws. By 2010, 48 percent of agricultural workers had zero to three years of schooling, as did 28 percent of construction workers, against only 13 percent in manufacturing. Agriculture and construction compete for the same workers; modern manufacturing, at least in Brazil, largely does not.

This pattern has broader relevance for development. Across many middle-income economies, including Brazil, manufacturing has become progressively more capital- and skill-intensive in recent decades, and the peak employment level the sector reaches along the path of development has declined—what Rodrik (2016) terms premature deindustrialization. A related observation, documented by Peters et al. (2026), is that many developing economies now transition from agriculture directly to non-tradable consumer services, bypassing the manufacturing stage altogether. The implication for structural transformation is that manufacturing expansion may no longer provide the ladder out of agriculture that it once did. Our evidence suggests that the expansion of construction—and potentially other low-skill service sectors—can fill that role instead, tightening agricultural labor markets and triggering reorganization.

The persistence of this labor-market tightening, and its lasting imprint on farm technology and structure, motivates the model we develop. Non-agricultural firms and farms are heterogeneous in productivity, with endogenous entry and exit along the lines of Hopenhayn (1992). The novel feature is on the agricultural side: each farm organizes production as a continuum of tasks ordered by their comparative advantage for capital versus labor (Zeira, 1998; Acemoglu and Restrepo, 2019), with the efficiency of labor relative to capital strictly decreasing in the task index, so that high-index tasks are naturally mechanizable and low-index tasks are not. The *mechanizable set*—the fraction of tasks the farm is technically equipped to perform with capital—is a state variable that expands only at a fixed adaptation cost I, making farm investments in making production tasks mechanizable irreversible—redrawing field plots, purchasing specialized equipment—.

The model yields three results. First, fixed costs reduce mechanization relative to the frictionless optimum, and larger farms mechanize more. Second, irreversibility implies that temporary wage shocks generate permanent mechanization: a farm that pays the adaptation cost when agricultural labor becomes relatively expensive does not revert to its pre-shock technology once wages fall back, and the resulting shift toward mechanized practices raises total factor productivity and induces land consolidation. Third, farm consolidation occurs whenever the non-agricultural wage rises faster than agricultural goods prices—a condition satisfied when agricultural goods face world prices, or when the non-agricultural shock does not fully pass through to agricultural output prices.

Two heterogeneity dimensions provide direct tests of this transmission mechanism. The first is informality: in microregions with higher baseline non-agricultural informality, the agricultural wage and mechanization gradients are flatter, consistent with Lewis (1954)'s characterization of the informal sector as a labor buffer that absorbs demand shocks before they reach agricultural wages. The second is terrain: the mechanization and agricultural-outcome gradients are largest at intermediate terrain steepness, not in the flattest microregions, which are typically already mechanized. Since our measure of mechanization is tractor adoption, the steepest areas face technical constraints regardless of factor prices; intermediate-terrain microregions—technically able to mechanize but not yet doing so—are instead where a shift in the cost of labor pushes farmers across their adoption threshold. Even where mechanization itself is attenuated, by high informality or steep terrain, the marginal farm exits once the relative cost of labor rises even modestly, consistent with farm selection.

Understanding agricultural reorganization matters for at least two reasons. First, the organization of agricultural production—farm size, capital intensity, and productivity—is a first-order determinant of agricultural output and food security in developing countries. An economy undergoing structural transformation is not simply one where agricultural employment shares fall, but one where the farms that remain active operate more productively. Second, the persistence of reorganization matters for long-run development. If farms that mechanize in response to a labor-demand shocks retain mechanization after the shock subsides, even transient episodes of non-agricultural growth can leave a lasting imprint on agricultural technology.

Our work connects four strands of literature, first of which is the literature on structural transformation (Gollin et al., 2007; Herrendorf et al., 2014; Alvarez-Cuadrado and Poschke, 2011; Rodrik, 2016; Bustos et al., 2019, 2016; Peters et al., 2026; Storesletten et al., 2026). Since the industrial revolution, the growth of the manufacturing sector has been praised to run in tandem with the economic development, followed, as economies grow richer, by the rise of the service economy. Accompanying this growth, there is a shrinking of the the labor force in the agricultural sector and improvements in productivity. Evidence from recent developers suggests that the role of the manufacturing sector as a stepping stone to the process of development might be upended, Rodrik (2016) and that the service economy may be absorbing labor released from agriculture, Peters et al. (2026). We add to this discussion the role of the *construction* sector, a sector with special features—relatively intensive in low-skill labor and non-tradable—that make it a natural outside option for agricultural workers at this stage of development

In the case of Brazil, Bustos et al. (2019) find that the workers released from agriculture were predominantly unskilled and that more-exposed regions industrialized but experienced slower manufacturing productivity growth. Much of the empirical work on Brazil exploits agricultural-technology shocks to identify the effect of *push* forces—labor released by agriculture—on structural change and agricultural productivity (Bustos et al., 2016). We study the complementary direction—a shock originating in non-agriculture—and, rather than stopping at the reallocation of labor, ask what happens to the farms that remain active, exploiting exploit quasi-random variation in the allocation of the largest low-income housing construction program in Latin-America to identify the effect of increases in non-agricultural labor demand onagricultural productivity and structural change.

Second, we contribute to the literature on agricultural organization, mechanization, and farm size (Adamopoulos and Restuccia, 2014; Manuelli and Seshadri, 2014; Caunedo and Kala, 2021; Caunedo and Keller, 2020; Boppart et al., 2023; Foster and Rosenzweig, 2022; Madhok et al., 2025; Asher and Novosad, 2020). Foster and Rosenzweig (2022) show that labor-market transaction costs explain why the smallest farms are more efficient than slightly larger ones in most low-income countries, while machine capacities that rise with operational scale generate the upper tail of productivity observed in rich-country agriculture; their accounting implies that consolidating farms toward minimum efficient scale would sharply raise output and income per farm worker. Closest to us is Madhok et al. (2025), who study agriculture's response when urbanization pulls labor out of rural India: there, migrant-sending households divest from farming and cultivate less land *without* substituting capital for labor or adopting machinery, and Asher and Novosad (2020) similarly find that rural roads move workers out of agriculture with little change in agricultural production. We document a different response in Brazil: a non-agricultural shock raises the relative cost of agricultural labor sufficiently, the farms that survive mechanize, consolidate, and raise output per worker through selection–moving up the size-productivity schedule that Foster and Rosenzweig (2022) characterize. The contrast indicates that agriculture's response to labor reallocation is not uniform but depends on whether the rising cost of labor crosses the threshold that makes capital deepening and farm exit profitable.

Third, we contribute to work on relative input prices, agricultural capital intensification, and the microeconomics of technology adoption (Manuelli and Seshadri, 2014; Boppart et al., 2023; Foster and Rosenzweig, 2010). The channel of relative-price changes driving capital intensification has been studied in the aggregate by Manuelli and Seshadri (2014) and Boppart et al. (2023). We focus on the micro-evidence, characterizing the anatomy of these shifts and the heterogeneity in responses that underlies the aggregates. This connects to the microeconomics of technology adoption surveyed by Foster and Rosenzweig (2010), who emphasize that adoption is the optimizing choice of heterogeneous agents facing heterogeneous returns, learning frictions, and credit and insurance constraints. We focus our analysis on the micro-evidence and characterize the anatomy of those shifts, as well as heterogeneity in responses driving those aggregates. We uncover a novel element of that process: even transitory shocks to relative input prices can induce farmers to incur the fixed cost of adoption, and because that investment is irreversible, a temporary shock generates persistent mechanization.

Finally, we add to research on the role of large employers in developing countries' local labor markets (Méndez and Van Patten, 2022; Felix, 2026; Sharma, 2025; Amodio et al., 2025). A large or dominant employer is a concentrated force in the local labor market, and its effect on wages depends on that market's structure—on employer concentration and wage-setting power (Felix, 2026; Amodio et al., 2025), and on the outside options available to workers, including self-employment and informal work, which attenuate employer market power (Amodio et al., 2025; Felix, 2026). Most of this work studies the incidence of such forces *within* the local non-agricultural labor market; our contribution is to follow them across the sectoral boundary, from the non-agricultural wages they shape to the reorganization of agriculture they induce. The transmission is itself shaped by local labor-market structure: consistent with Lewis (1954)'s account of the informal sector as a labor reservoir, the agricultural response is attenuated where non-agricultural informality is high (a margin that is quantitatively large in Brazil, Ulyssea, 2018), because the demand shock is absorbed in the reallocation from informal to formal employment before it reaches agricultural wages.

2 Non-Agricultural Labor Demand Farm Reorganization

This section documents two sets of reduced-form results on whether and how non-agricultural labor demand drives structural change in Brazil. The first exploits the staggered entry of large non-agricultural firms across local labor markets to identify the impact of a labor-demand shock on formal-sector employment and wages. The second asks whether a large federal housing-construction program—Minha Casa Minha Vida (MCMV)—produces reorganization of agricultural production. The two designs are complementary: the large-entry quasi-experiment allows high-frequency annual event-study estimation, but the agricultural outcomes we ultimately care about are available only for three census rounds; the MCMV design provides a second source of variation specifically in the labor demand of construction sector, and its spatial allocation being partially driven by political-economy predictors that are plausibly orthogonal to local agricultural conditions.

2.1 Data and Sample Construction

This section describes the data sources we use the estimate the effects of new large firm entry on the local labor markets and agricultural sector, and to estimate our model. Our main data sources are the Brazilian agricultural census and linked employer-employee confidential dataset Relação Anual de Informações Sociais (RAIS).[1] The datasets used in the analysis differ in their frequency and coverage. Annual data from RAIS provide a high-frequency picture of the formal labor market from 1986 to 2016, while agricultural outcomes are available only at decadal intervals from the Agricultural Census (1997, 2006, 2017). The population census (1991, 2000, 2010) provides additional labor market variables, including informality and migration.[2] Additional data details, including coverage timeline of every dataset and the two treatment windows (firm entry, 1998–2017; MCMV, 2009–2014; SICONV, 2008–2017), are available in Appendix B.

Geographic Units. The analysis is conducted at the Minimum Comparable Area (MCA) – a time consistent aggregation of municipalities, as well as at the microregion level (MMC). MCAs are the smallest consistent sets of municipalities. For ease, we use the term municipality to refer to MCAs. There are 5,352 MCAs, covering 5,507 municipalities, between 1997 and 2017. We construct MCAs using information from the Brazilian Census on which municipalities split and/or emerged as combinations of which municipalities between 1991 and 2010.

Agricultural Sector. We use the Brazilian agricultural census data for years 1995-96, 2006, and 2017 to obtain information on the organization of the agricultural sector. We study farm area, agricultural labor,

[1] We use the RAIS data cleaned by Engbom and Moser (2022).

[2] We build measures of internal migration across municipalities from the population census, disaggregated by socio-economic characteristics. Preliminary analyses, not included in this draft, suggest large entries draw working-age male migrants into epicenter in the 5-year window during entry.

value of produced output, and number of tractors employed. We also obtain information on input expenditure shares as well as on the number of agricultural establishments (farms). We further classify these establishments by land size: a small farm if less than 100 hectares (ha henceforth), a medium farm if between 100 and 500 ha, and a large farm if holding more than 500 ha. In addition, we obtain land usage rights to determine how much of used land is rented and how much is owned. We also use the Pesquisa Agrícola Municipal (PAM), an annual municipal-level survey of crop production published by IBGE and available from 1974 to 2019. PAM provides the quantity, harvested area, and production value for the main temporary and permanent crops at the municipal level. We use PAM to construct agricultural output per worker at the microregion level and to study crop-level production responses—corn, soy, and sugarcane.

Labor Markets. We use three complementary data sources to derive labor market outcomes: RAIS, which covers nearly the universe of formal employees on an annual basis between 1986 and 2017; the Agricultural Census for agricultural labor, available for 1997, 2006, and 2017; and the Brazilian Population Census, available for 1991, 2000, and 2010 for information on informal employment and internal migration.

From RAIS we identify and describe in detail the worker characteristics of all large firm entries, obtain municipality-level data on formal employment, earnings, wages (earnings per contract hours), separately by economic activity sector and by worker education group.

The Agricultural Census (Ag Census henceforth), provides information on the total agricultural labor at the municipality level. RAIS also provides information for agricultural labor, but only includes formal employment. Our empirical analysis includes annual estimates using data from RAIS, as well as decadal estimates for the outcomes from Agricultural Census. To match the timing of Agricultural Census, we use annual data or data for the entire decade preceding years 1997, 2006 and 2016.

We use the population Census to document broader patterns of employment and earnings, including formality rates, separately by sectors and education groups. We restrict our Census analyses to individuals aged 18-64. Informal workers are identified as those without a signed *Carteira de Trabalho* (employee registration), as is standard in the literature (Ulyssea, 2018).

Minha Casa Minha Vida. We use data of the spatial allocation of MCMV projects, which depended on a combination of federal targeting criteria—most notably measures of local housing deficits—developer proposals satisfying program rules, and the availability of government-owned land at the municipal, state, or federal level. As documented by Simoni Junior and Dias (2025), mayoral political alignment with the federal ruling coalition in 2009 strongly predicts the allocation of MCMV contracts for the lowest-income housing tier during 2009–2014, prior to the 2010 presidential election. These institutional features generate plausibly exogenous variation in local construction activity. We measure MCMV intensity as units contracted over 2009–2014 per 100 baseline dwellings. We complement MCMV with a second source of federally financed construction investment: the SICONV transfer system (Sistema de Gestão de Convênios e Contratos de Repasse), which records voluntary (discretionary) intergovernmental transfers from federal agencies to municipalities for projects including urban infrastructure, road paving, and rural development. We measure SICONV intensity as committed *repasse* per capita over 2008–2017. Because SICONV transfers are discretionary, political-economy predictors are informative about their cross-municipality allocation. Additional institutional details on both programs are provided in Appendix C.

2.2 Motivating Facts

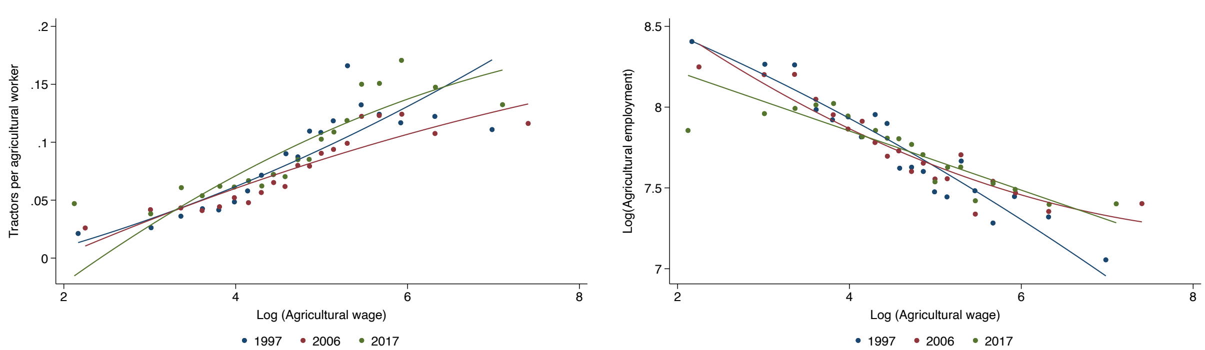
Before turning to our empirical strategy, we document a set of stylized facts that motivate our analysis and contextualize our empirical approach. These facts, drawn from national and municipal-level data, illustrate the ongoing process of structural transformation in Brazil and provide suggestive evidence that localized labor demand shocks originating in non-agriculture can affect agricultural outcomes.

Brazil offers an ideal setting for this analysis: Brazil's continental size and regional diversity generate substantial heterogeneity in the level of income, pace and character of structural transformation across municipalities, providing rich cross-sectional and time-series variation to exploit.

Figure 1: Organization of Agricultural Production and Agricultural Wage

Panel A: Tractors per worker

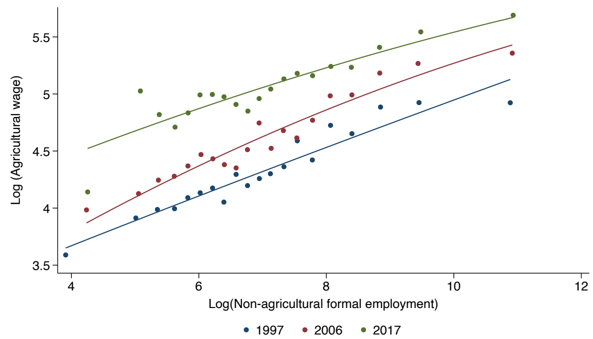
Panel B: Agricultural employment



Notes: This figure plots binned scatters showing cross-municipality variation, separately by year and conditional on region fixed effects, for outcomes measured at municipality level and aggregated by MCAs (Minimum Comparable Areas). Data on tractors, agricultural wages, and agricultural employment are from the Agricultural Censuses for 1997, 2006, and 2017. Agricultural wages are calculated as the sum of nominal labor expenses divided by the total number of workers employed in agriculture in each municipality. Wages are converted to 2017 reais using the IPCA inflation index.

At the municipal level, we can correlate information on the cost of labor, i.e. wages, and agricultural mechanization. Higher agricultural wages are associated with greater mechanization, lower agricultural employment and higher formal non-agricultural employment. Figure 1 presents binned scatter plots using municipality-level data, aggregated to Minimal Comparable Areas (MCA) for cross-time comparisons, for each Agricultural Census round. Panel A shows that tractors per agricultural worker is increasing in the agricultural wage, while Panel B shows that agricultural employment is decreasing in the same measure. That is, municipalities where labor is relatively more expensive exhibit more capital-intensive

Figure 2: Agricultural Wage and Employment in Non-Agriculture



Notes: This figure plots binned scatters showing cross-municipality variation, separately by year and conditional on region fixed effects, outcomes measured at municipality level and aggregated by MCAs (Minimum Comparable Areas). Data on agricultural wages are from the Agricultural Censuses for 1997, 2006, and 2017, converted to 2017 reais using the IPCA inflation index. Data for employment are for the formal sector only, from RAIS for the respective years. Agricultural wages are calculated as the sum of nominal labor expenses divided by the total number of workers employed in agriculture in each municipality.

farms and lower agricultural employment. The three census-round series essentially overlap, indicating a stable cross-sectional relationship. Figure 2 shows that agricultural wages in a municipality increase in the level of formal non-agricultural employment. A plausible interpretation is that tighter labor markets in non-agriculture—or higher non-agricultural productivity—spill over into agriculture, putting upward pressure on rural wages. Hence, labor-demand shocks originating in non-agriculture can affect agricultural production decisions through general-equilibrium wage effects.

Taken together, these patterns motivate our empirical approach. First, structural transformation is ongoing in Brazil over our study period, with agricultural employment declining and production becoming increasingly capital-intensive. Second, this process is heterogeneous across regions, and local labor market conditions —as summarized by wages in the agriculture and non-agricultural sector— correlate with the relative employment shares across sectors and importantly, with capital intensity.

Since these are equilibrium outcomes in the labor and output markets, it is not possible to interpret these wage differentials, and associated outcomes, as causal. To learn more about the role of wage increases on the organization of production in the agriculture sector, we study large firm entries into local labor markets and build a staggered difference-in-differences research design. We complement this analysis with a second empirical strategy centered specifically on the construction sector. Two considerations motivate this focus. First, the large-entry results below show that effects on wages and employment concentrate in construction and real estate entries, with manufacturing entries generating smaller and less precise estimates.

These patterns are consistent with modern manufacturing in Brazil having become increasingly capital-intensive and skill-intensive: Table 1 shows, using Brazilian Demographic Census data for 1991–2010, that 45% of manufacturing workers completed secondary education by 2010, compared with only 20% in

Table 1: Skill and Demographic Composition of 21st Century Employment in Brazil

	1991			2000			2010		
	Agric.	Constr.	Manuf.	Agric.	Constr.	Manuf.	Agric.	Constr.	Manuf.
Share of employment (%)	21.2	7.0	15.2	16.7	7.2	13.5	12.9	7.5	12.2
Skill (years of schooling)									
Mean years of schooling	2.48	4.19	6.34	3.19	5.21	7.25	4.38	6.13	8.46
% with 0–3 years	66.2	42.5	21.2	55.8	30.3	15.1	48.0	28.0	13.2
% with 4–7 years	26.8	40.8	40.2	32.1	43.6	34.5	27.6	31.5	19.7
% with 8–10 years	3.7	8.8	17.7	6.1	13.8	21.4	13.2	20.0	22.2
% with 11–14 years	2.6	5.4	15.8	3.6	9.1	23.4	9.7	17.0	37.3
% with 15+ years	0.7	2.4	5.0	0.6	2.4	4.7	1.5	3.5	7.6
% completed secondary ($\geq$11 yrs)	3.3	7.9	20.8	4.3	11.5	28.1	11.2	20.5	44.9
Demographics									
% female	12.5	4.4	25.6	22.1	3.6	31.6	29.0	3.6	34.9
% White	43.4	45.4	60.8	44.6	47.5	61.8	39.8	39.3	54.7
% Brown (pardo)	49.5	46.5	33.3	45.9	42.3	31.2	50.1	47.9	36.7
% Black (preto)	6.3	7.5	5.0	7.9	9.1	5.7	8.2	11.6	7.4
% Asian	0.3	0.2	0.6	0.3	0.2	0.5	1.0	0.9	1.0
% Indigenous	0.4	0.1	0.1	0.7	0.4	0.3	0.9	0.3	0.2
Sample size									
Weighted employment (millions)	10.46	3.48	7.49	10.06	4.35	8.14	10.34	6.00	9.81
Person records (unweighted)	1,437,582	379,652	794,652	1,516,609	491,466	905,935	1,696,114	616,376	1,013,767

Notes: Person-level IBGE Demographic Census microdata, employed workers aged 18–64, weighted by `xweighti`. "Share of employment" is each sector's weighted share of all employed workers aged 18–64 in that wave. Sectors are defined from `atividade` (1991) and `cnae` (2000/2010); extractive/mining is excluded. "% completed secondary" is the weighted share with 11 or more years of schooling. Race: White, Brown (*parda*), Black (*preta*), Asian, Indigenous. Education and race shares need not sum exactly to 100 because of records with missing schooling or race.

construction and 11% in agriculture. In contrast, agriculture and construction share a similar skill profile at the low end—in 2010, 48% of agricultural workers and 28% of construction workers had 0–3 years of schooling, versus 13% in manufacturing. This skill overlap implies that a construction boom poses a tighter outside option for agricultural workers than a manufacturing boom does.

2.3 Large Firm Entry and Local Labor Markets

This section documents the effects of large non-agricultural firm entries on formal-sector labor market outcomes. We first describe the large entry shocks, their geographic distribution, and their sectoral composition. We then present the staggered difference-in-differences design and the main results, focusing on how entries affect formal employment and wages and how effects vary by sector and worker education.

2.3.1 Description of the "Shock"

Our goal is to identify non-agricultural firm entry events that are large relative to the local labor market and therefore generate a shift in labor demand, bidding up local wages.

We define a large entry shock as a non-agricultural new establishment entry comprising more than 0.1% of its microregion's population in the year preceding the shock and with at least 200 employees at entry, roughly the median entry size of all non-agricultural entries that are at least 0.1% of the microreigon population (Figure A1).[3] The former restriction identifies new entries that are large enough relative to a potential labor market of a region, whereas the second removes very small regions, for which even a 5-employee entry (5th percentile of new entering establishments) would otherwise be considered a large entry. Another concern with interpreting the effects of these large entry shocks as exogenous to the agriculture sector is that the firms entering these local labor markets may also rely heavily on agricultural goods as inputs, and, hence, choose the location based on inputs availability and proximity. To address this, in our baseline estimates we exclude large entries in sectors directly demanding agricultural products, such us Food and Beverages, Textiles, and Rubber and Tobacco.

In addition, given that we are interested in shocks that may draw workers away from agriculture, we exclude high-skilled sectors whose labor pool overlaps little with the agricultural workforce, specifically public administration, education, finance, and healthcare. Entries are classified into three mutually exclusive sectors: construction and real estate, manufacturing, and other (comprising transport and communication, hospitality, mining, and wholesale trade). The analysis restricts to entries occurring between 1998 and 2017, so that the 1997 Agricultural Census serves as a clean pre-entry baseline; microregions with any qualifying large entry before 1998 are dropped.

Table A1 summarizes the analyzed entries. Of 81 epicenter municipalities, 57 host a construction or real estate entry, 15 a manufacturing entry, and 9 an "other" entry. Construction and real estate entries are larger on average (mean of 502 employees at entry) and draw a notably larger share of primary-education workers (18% of entrants) than manufacturing (13%), with most workers coming from other formal (RAIS) jobs. Figure A1 shows the size distribution of all new firm entries with employment at entry above 0.1% of its microregion population.

2.3.2 Empirical Design: Staggered Difference-in-Differences

We classify municipalities into three groups. Municipalities that experienced at least one large entry shock between 1998 and 2017 are defined as *epicenters*. Municipalities that never experienced a large entry but are located within the same micro-region as an epicenter are defined as *surrounding* municipalities. Finally, municipalities that never experienced a large entry and are located in micro-regions without any epicenters are classified as *never-entered* ones. Appendix B.5 details the precise construction of these groups.

To estimate the effect of large entries on formal-sector outcomes—primarily in non-agriculture—we compare *epicenters* to municipalities in *never-entered* microregions.[4] A strength of this comparison is that control municipalities are never affected by large entries. The high frequency of the data and the large number of municipalities lets us implement with precision a staggered difference-in-differences estimator

[3]Combined, these restrictions ensure that the entries we focus on are not only large relative to the local economy, but also that they are entries in large enough economies. Relaxing either constraint results in the inclusion of municipalities like the island of Fernando de Noronha, which is a natural reserve and would otherwise be the municipality with the largest number of "large entries", 28, per the micro-region population cutoff in the data.

[4]While most of the formal sector in Brazil is non-agricultural, we also apply this design to formal agricultural employment to capture some of the effects at higher frequency.

robust to heterogeneous treatment effects, à la De Chaisemartin and d'Haultfoeuille (2020):

$$\widehat{DID}_{\ell} = E\big[\big(y_{i,t_i+\ell} - y_{i,t_i-1}\big) - \big(y_{i',t_i+\ell} - y_{i',t_i-1}\big)\big]\,, \quad \ell = -5, \ldots, +10, \tag{1}$$

where i denotes an epicenter municipality, t_i its entry year, and i' indexes never-entered control municipalities. Each $\widehat{DID}_{\ell}$ is the cumulative effect of ℓ periods of exposure, measured relative to the year before entry ($t_i - 1$); estimates for $\ell < 0$ serve as a test of parallel pre-trends.[5] We report pooled epicenter-vs-never effects as well as effects with the treated side restricted to a single entry sector (construction and real estate, manufacturing, or other), each against the same never-entered control group.

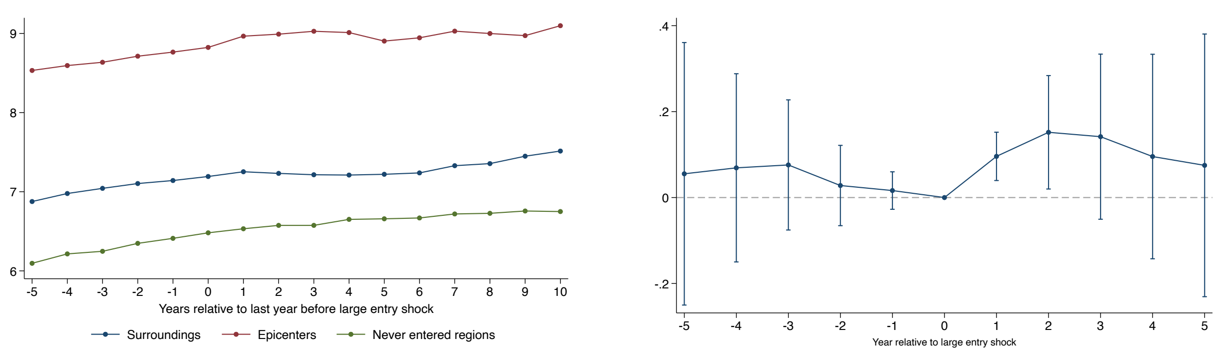
2.3.3 Results

The large-entry results reveal an effect that is both sectorally concentrated and skill-specific—two features that link the formal labor-market findings directly to the agricultural reorganization documented in the next section. Figure 3 shows that a large non-agricultural firm enters a municipality, formal employment at the epicenter rises by about 10 log points relative to never-entered controls in the year of entry, with parallel trends between epicenter and never-entered municipalities prior to entry.

Figure 3: Log formal non-agricultural employment

Panel A: Average levels by group

Panel B: Epicenter vs. never-entered DD



Notes: Panel A shows average outcome levels by municipality type. Panel B shows cumulative effects of large entries at epicenters, relative to never-entered controls, per De Chaisemartin and d'Haultfoeuille (2020)'s staggered difference-in-differences estimator. Never-entered regions are assigned an random entry year for Panel A. Standard errors are clustered at the microregion level.

Construction is the operative sector. But this average masks a sharp sectoral asymmetry: the effect is driven entirely by construction and real estate entries, which generate a 13 log-point increase in formal non-

[5] We estimate the effects using Stata's `did_multiplegt_dyn` package, allowing for group-specific linear trends, using never-entered municipalities as controls, and clustering standard errors at the microregion level. Group-specific linear trends estimates the event study on the first difference of the outcome, allowing each municipality its own linear trend, so identification requires treated and never-entered units to deviate from their group-specific linear trends in parallel rather than to share parallel outcome levels. The pre-period placebo coefficients then test this assumption—flat, statistically-zero placebos indicate no differential pre-trend beyond a common linear drift.

agricultural employment at year one (Table 2). Manufacturing and other-sector entries produce coefficients near zero that are indistinguishable from the pre-entry trends (Figures A3 and A6).

Table 2: Cumulative effect of large entry shocks on formal employment at epicenters, by entry sector

	Construction / real estate			Manufacturing			Other		
Outcome	Yr 1	Yr 5	Yr 10	Yr 1	Yr 5	Yr 10	Yr 1	Yr 5	Yr 10
Entry-shock size (formal employees)	544.347***	0.000	0.000	376.894***	0.000	0.000	378.735***	0.000	0.000
	(70.273)	(0.000)	(0.000)	(66.784)	(0.000)	(0.000)	(110.317)	(0.000)	(0.000)
Log formal employment (total)	0.112***	0.067	0.320	0.034	0.130	0.277	-0.005	-0.167	-0.167
	(0.037)	(0.201)	(0.626)	(0.031)	(0.148)	(0.393)	(0.034)	(0.241)	(0.852)
Log formal employment (non-ag)	0.132***	0.118	0.522	0.038	0.154	0.302	-0.024	-0.288	-0.479
	(0.042)	(0.224)	(0.691)	(0.034)	(0.159)	(0.425)	(0.032)	(0.219)	(0.696)
Log formal employment (ag)	-0.015	-0.184	-0.541	0.002	-0.014	0.153	0.093	0.576	0.451
	(0.047)	(0.164)	(0.483)	(0.060)	(0.238)	(0.269)	(0.104)	(0.472)	(1.680)

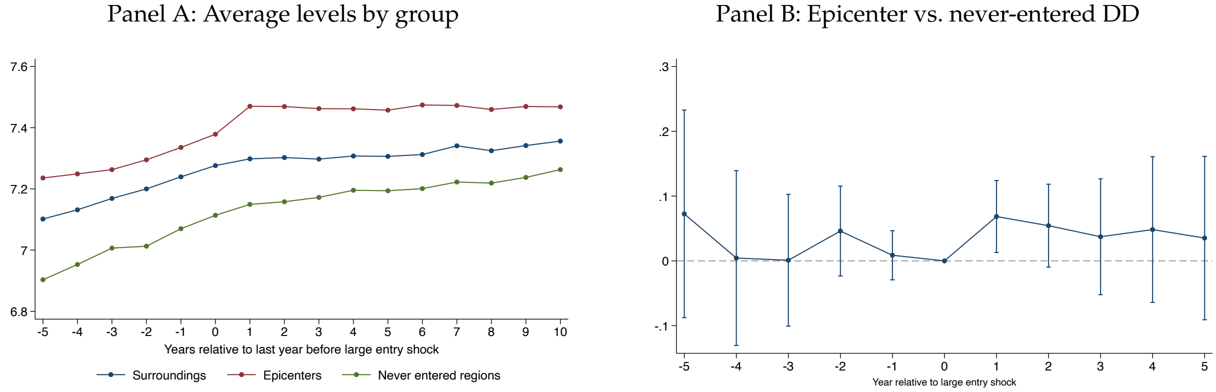
Notes: This table reports cumulative effects of large entries at epicenters, relative to never-entered controls, per De Chaisemartin and d'Haultfoeuille (2020)'s staggered difference-in-differences estimator, at years 1, 5, and 10 after entry. Standard errors are clustered at the microregion level. Each column group restricts the treated set of municipalities (epicenters) to a single entry sector, whose effects are estimated against the same never-treated control group. Standard errors clustered at the microregion level in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

The contrast traces directly to skill composition. By 2010, construction workers and agricultural workers shared a predominantly primary-educated labor force—28% of construction workers had 0–3 years of schooling, as did 48% of agricultural workers, versus only 13% in manufacturing (Table 1). A construction firm therefore competes for the same workers as local farms; a manufacturing firm draws from a different pool and does not. Construction entries temporarily raise formal employment for primary education workers, manufacturing entries have no short-run or long-run effect for that group (Table 3, Figure A4).

On average, large entries raise average real monthly earnings by about 7 log points at entry (Figure 4), with these effects also being driven by construction entries, especially for workers with at most primary education (Table 3, Figure A5).

The outside option for agricultural workers rises sharply—and persistently. Real wages for primary-education workers—those who most directly overlap with the agricultural labor force—rise by 12 log points at year one and continue to grow, reaching 52 log points above the never-entered baseline by year 10. This is not a transitory local labor-market shock: a decade after entry, the relative return to staying in agriculture has permanently deteriorated for the workers who constitute the agricultural workforce. At the same time, the shock to formal employment is transitory.

Figure 4: Log non-agricultural average real log monthly earnings



Notes: Panel A shows average outcome levels by municipality type. Panel B shows cumulative effects of large entries at epicenters, relative to never-entered controls, per De Chaisemartin and d'Haultfoeuille (2020)'s staggered difference-in-differences estimator. Standard errors are clustered at the microregion level.

Table 3: Heterogeneous cumulative effects of large entry shocks on formal employment and earnings, by firm entry sector and worker education group

	Construction / real estate			Manufacturing			Other		
Education group	Yr 1	Yr 5	Yr 10	Yr 1	Yr 5	Yr 10	Yr 1	Yr 5	Yr 10
Panel A. Log formal employment									
Primary education	0.147***	-0.068	0.206	0.028	0.128	0.219	0.032	0.104	0.246
	(0.045)	(0.213)	(0.624)	(0.035)	(0.208)	(0.485)	(0.039)	(0.219)	(0.781)
Secondary education	0.117***	0.075	0.346	0.043	0.163	0.240	-0.023	-0.300	-0.586
	(0.039)	(0.202)	(0.633)	(0.034)	(0.137)	(0.349)	(0.035)	(0.264)	(0.945)
Tertiary education	0.065	0.142	0.276	-0.010	-0.116	0.273	-0.020	-0.171	-0.257
	(0.046)	(0.191)	(0.628)	(0.034)	(0.165)	(0.372)	(0.048)	(0.291)	(0.933)
Panel B. Log average real per-hour wage									
Primary education	0.123***	0.126	0.517***	0.037	0.065	0.380	0.074	-0.309	-0.591
	(0.037)	(0.105)	(0.196)	(0.036)	(0.083)	(0.258)	(0.127)	(0.249)	(0.785)
Secondary education	0.076**	0.101	0.256	0.029*	0.028	0.248***	0.101	-0.103	-0.258
	(0.034)	(0.099)	(0.241)	(0.015)	(0.027)	(0.080)	(0.129)	(0.125)	(0.437)
Tertiary education	0.069	0.032	-0.131	0.028	-0.059	-0.294	-0.030	-0.365	-0.826
	(0.044)	(0.109)	(0.299)	(0.039)	(0.167)	(0.638)	(0.049)	(0.261)	(0.749)

Notes: This table reports cumulative effects of large entries at epicenters, relative to never-entered controls, per De Chaisemartin and d'Haultfoeuille (2020)'s staggered difference-in-differences estimator, at years 1, 5, and 10 after entry. Standard errors are clustered at the microregion level. Each column group restricts the treated set of municipalities (epicenters) to a single entry sector, whose effects are estimated against the same never-treated control group. Standard errors clustered at the microregion level in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

The gain for secondary-education workers is positive but smaller; the effect on tertiary wages is indistinguishable from zero, consistent with negligible overlap between highly educated workers and agricultural employment.

Labor moves out of agriculture, not into it. Formal agricultural employment at construction epicenters is a precise zero in year one and turns negative by years 5–10 (Figure A6). The employment pull runs in one direction: construction expands while agriculture contracts. This directional pattern—lower formal agricultural employment alongside higher non-agricultural wages for less-educated workers—is the hallmark of a pull-driven structural transformation rather than an agricultural demand expansion or a skill-neutral productivity shock.

Having established that construction entry tightens the outside option for the agricultural labor force, we now ask whether a policy-driven construction shock—the Minha Casa Minha Vida program—produces the downstream reorganization of agricultural production that this wage-and-labor story predicts.

2.4 Housing Construction and Agricultural Reorganization

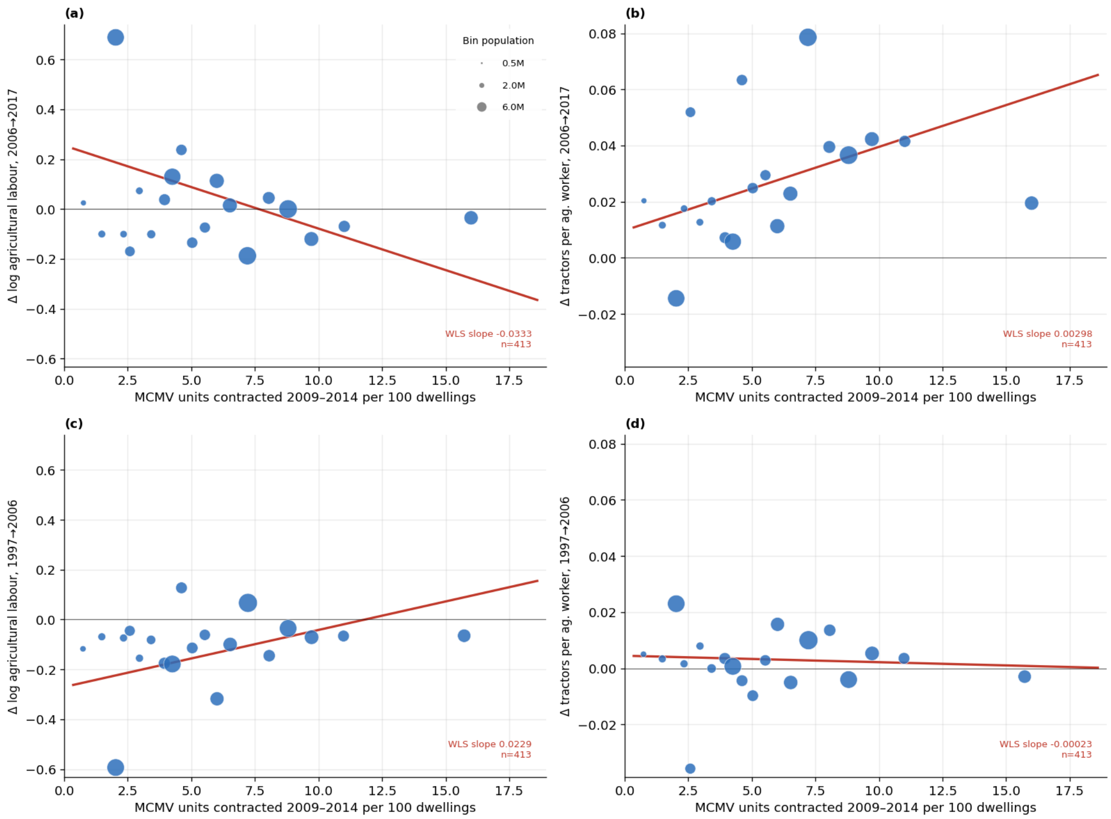
The large-entry evidence identifies construction and real estate entries as the dominant driver of formal employment and wage gains among less-educated workers. We now ask whether a policy-induced source of construction labor demand—the federal Minha Casa Minha Vida (MCMV) housing program—generates the same reorganization of agricultural production. We pursue this strategy for three reasons. First, construction entries are the main source of labor-market effects documented above, making construction the natural sector through which to test the agricultural reorganization mechanism. Second, the agricultural census is available for only three rounds (1997, 2006, 2017), making a dynamic event-study for agricultural outcomes infeasible; the MCMV design, which exploits cross-municipal variation in program intensity, is suited to detecting decadal changes. Third, the spatial allocation of MCMV contracts was partially driven by political-economy factors—in particular, mayoral alignment with the federal governing coalition—that are plausibly orthogonal to local agricultural conditions.

2.4.1 MCMV as a Source of Construction Demand

Before presenting the main MCMV results, we document two things. First, MCMV intensity predicts large construction entries. Table A2 shows that one additional MCMV unit per 100 dwellings raises the probability of a construction or real estate large entry by approximately half a percentage point, while the coefficient for other-sector entries is small and insignificant; SICONV housing transfers do not predict entries. This establishes that the MCMV mechanism operates through direct construction activity rather than general municipal fiscal capacity.

Second, Figure 5 plots the observed relationship between MCMV intensity and the changes in agricultural labor and tractors per worker over 2006–2017, together with a 1997–2006 placebo. Higher MCMV intensity is associated with lower agricultural employment and more tractors per worker during the treatment period, with no similar pattern in the placebo period—a first suggestive piece of evidence that housing construction drives agricultural reorganization.

Figure 5: Observed MCMV intensity vs. AgCensus change: agricultural labour and tractors per worker.



Notes: This figure shows binned scatters of MCMV's housing program intensity at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and decadal changes in agricultural organization outcomes, measured by the 1997, 2006, and 2017 wages of the Agricultural census. Δ log agricultural labour (left) and Δ tractors per worker (right), over 2006$\rightarrow$2017 (top) and the 1997$\rightarrow$2006 placebo (bottom). Binned means are population-weighted and the red line shows a weighted-least squares fit over $x \leq$99th percentiles; marker area is proportional to bin population.

2.4.2 Design: Program Intensity and Allocation Predictors

Because MCMV units are not allocated randomly, we isolate the component of MCMV intensity driven by political-economy allocation predictors, following Simoni Junior and Dias (2025), whose main results we first replicate in Table A4. Let MCMV_m denote contracted units per 100 dwellings in microregion m. We estimate the first-stage regression

$$\text{MCMV}_m = \text{KEY}'_m\beta + X'_m\gamma + \delta_{r(m)} + u_m, \tag{2}$$

where KEY_m collects four political-economy predictors: an indicator for a mayor in the federal governing coalition in 2009, the core and swing PT presidential vote shares (1994–2006), and a PT state-governor indicator (2006). X_m are 2000-vintage socioeconomic controls (urban population share, a below-50,000-inhabitants indicator, a 2005 metropolitan-region indicator, and the agricultural, manufacturing, and formal employment shares), and $\delta_{r(m)}$ are macro-region fixed effects. The predicted intensity is $\widehat{\text{MCMV}}_m = \text{KEY}'_m\hat{\beta}$, the fitted contribution of the political predictors only. We then relate predicted intensity to the predicted change in each outcome, $\widehat{\Delta y}_m$, constructed analogously:

$$\widehat{\Delta y}_m = \theta\,\widehat{\text{MCMV}}_m + \varepsilon_m. \tag{3}$$

Equation (3) is a population-weighted bivariate regression whose slope θ summarizes how the politically driven component of housing intensity co-moves with each outcome. The first-stage allocation results (Table A3) confirm that mayoral alignment and metropolitan-region status are the strongest predictors of MCMV allocation, while the measured housing deficit is not, motivating the predicted-on-predicted design. Because the first stage is not strong enough to support instrumental variables and an exclusion restriction is implausible, θ is a descriptive association rather than a causal estimate, to be read alongside the large-entry quasi-experiment. We report results for a second source of construction investment—SICONV federal infrastructure transfers—using the same predicted-on-predicted design in Appendix C.2.

2.4.3 Results

Annual outcomes confirm the link to the large-entry evidence. The predicted-MCMV gradient in annual RAIS data (Table 4) replicates the large-entry pattern at the program level. A one-standard-deviation increase in predicted MCMV intensity is associated with roughly 9–10 log points faster growth in formal construction employment over 2009–2017 (0.095, column 1), while formal agricultural employment falls over the same period (Appendix Figure A7). The sign and direction of these annual RAIS gradients match the large-entry event-study findings: construction demand rises, agricultural employment contracts. This consistency across two independent sources of identification—firm entry and housing-program intensity—is the first piece of evidence that the MCMV mechanism operates through the same labor-market channel.

Agriculture reorganizes: lower employment, higher wages, more capital, bigger farms. and Tables 5–6 and Figures 6–8 present the decadal agricultural census results for 2006–2017. Reading the single-instrument MCMV column, a one-standard-deviation increase in predicted MCMV intensity is associated with agricultural employment falling by about 1 log point (0.011), agricultural wages rising by about 2 log points (0.018), tractors per worker rising by 0.006 units, average farm size rising by about 6 log points (0.060), and agricultural output per worker rising by about 17 log points (0.167)—and every one of these coefficients is statistically significant.

Table 4: Predicted MCMV and SICONV intensity and 2009→2017 RAIS outcomes

	Predicted outcome change (2009→2017)			
	(1) MCMV	(2) SICONV	(3) Both	(4) Net
Panel A. Predicted Log Construction Employment				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.0946***		0.147***	
	(0.018)		(0.016)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.000218	-0.074***	
		(0.014)	(0.016)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.066***	
			(0.014)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.0689***
				(0.017)
Pre-period actual change (control)	0.0462***	0.0482***	0.0398***	
	(0.016)	(0.015)	(0.012)	
N (microregions)	381	381	381	381
Panel B. Predicted Log Primary-Education Wage				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.00437		0.0045	
	(0.0031)		(0.0034)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.00569**	0.00168	
		(0.0028)	(0.0038)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.0109***	
			(0.0029)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.00656*
				(0.0035)
Pre-period actual change (control)	0.0707***	0.0812***	0.0677***	
	(0.01)	(0.012)	(0.011)	
N (microregions)	394	394	394	394

Notes: Predicted intensities are standardized, so each slope is the effect of a 1 SD increase and the MCMV and SICONV coefficients are comparable. Regressions are weighted by 2000 population, trimmed to the 1st–99th percentile of both intensities. Standard errors are clustered at the microregion; every column controls for pre-treatment (2001-2009) outcome changes. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Crucially, all five outcomes in Tables 5–6 and Figures 6–8 move in the same direction simultaneously, and a placebo test using the pre-program period (1997–2006) shows a flat gradient on mechanization and, if anything, a slight *positive* gradient on agricultural employment (Figure 5), ruling out pre-existing trends as the source of the pattern.

The five-way joint pattern is the mechanism at work: a construction-driven increase in the cost of agricultural labor makes labor-intensive farming less profitable, inducing a substitution of capital for labor, the exit of the least productive farms, and the consolidation of land in surviving operations. These surviving farms are more productive precisely because the exit of the marginal farms has raised the average.

Table 5: Predicted MCMV and SICONV intensity and predicted 2006→2017 Agricultural labor

	Predicted outcome change, treatment period (2006→2017)			
	(1) MCMV	(2) SICONV	(3) Both	(4) Net
Panel A. Predicted Log Agricultural Wage				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.0178**		-0.0344***	
	(0.007)		(0.0038)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.0764***	0.0913***	
		(0.0052)	(0.0037)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.0153***	
			(0.0032)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.0564***
				(0.0039)
Pre-period actual change (control)	-0.00479	-0.0025	0.000481	
	(0.015)	(0.0071)	(0.004)	
N (microregions)	394	394	394	394
Panel B. Predicted Log Agricultural Employment				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	-0.0111**		-0.0122**	
	(0.0053)		(0.0055)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		-0.0102**	-0.00243	
		(0.0043)	(0.0054)	
Predicted MCMV × Predicted SICONV (SD × SD)			-0.0227***	
			(0.0047)	
Net effect (1 SD MCMV + 1 SD SICONV)				-0.0139**
				(0.0058)
Pre-period actual change (control)	0.00455	0.000433	0.0086	
	(0.011)	(0.012)	(0.0087)	
N (microregions)	394	394	394	394

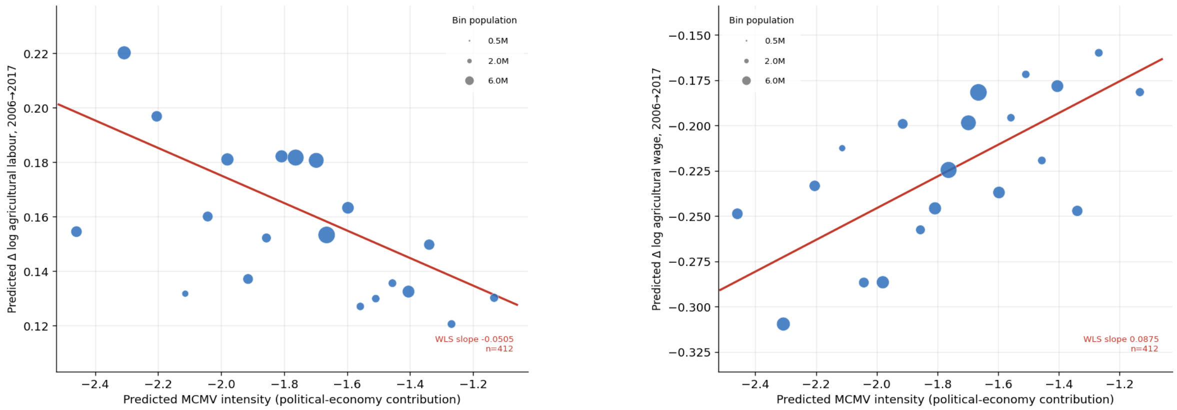
Notes: Outcomes measured per the Agricultural Census. See notes to Table 4.

Table 6: Predicted MCMV and SICONV intensity and 2006→2017 mechanization

	Predicted outcome change, treatment period (2006→2017)			
	(1) MCMV	(2) SICONV	(3) Both	(4) Net
Panel A. Predicted Tractors per Worker				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.00606***		0.00481***	
	(0.00025)		(0.0002)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.00443***	0.00216***	
		(0.00047)	(0.00033)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.000178	
			(0.00024)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.00697***
				(0.00028)
Pre-period actual change (control)	0.0027	-0.00187	0.00247	
	(0.0084)	(0.0074)	(0.0064)	
N (microregions)	394	394	394	394
Panel B. Predicted Log Farm Size (ha/establishment)				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.0603***		0.0609***	
	(0.0015)		(0.0016)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.0278***	-0.000924	
		(0.0054)	(0.0024)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.00117	
			(0.0019)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.06***
				(0.002)
Pre-period actual change (control)	0.0102	0.00694	0.0103	
	(0.008)	(0.011)	(0.008)	
N (microregions)	394	394	394	394
Panel C. Predicted Log Agricultural Output per Worker				
Predicted MCMV intensity (per 1 SD, units/100 dwellings)	0.167***		0.198***	
	(0.0066)		(0.0051)	
Predicted SICONV intensity (per 1 SD, R$/1,000 residents)		0.0432***	-0.0519***	
		(0.016)	(0.0047)	
Predicted MCMV × Predicted SICONV (SD × SD)			0.0166***	
			(0.0051)	
Net effect (1 SD MCMV + 1 SD SICONV)				0.146***
				(0.0053)
Pre-period actual change (control)	0.00746	-0.016	0.00886	
	(0.0088)	(0.022)	(0.0069)	
N (microregions)	389	389	389	389

Notes: Outcomes measured per the Agricultural Census. See notes to Table 4.

Figure 6: Predicted AgCensus labor and wage changes vs. predicted MCMV intensity.

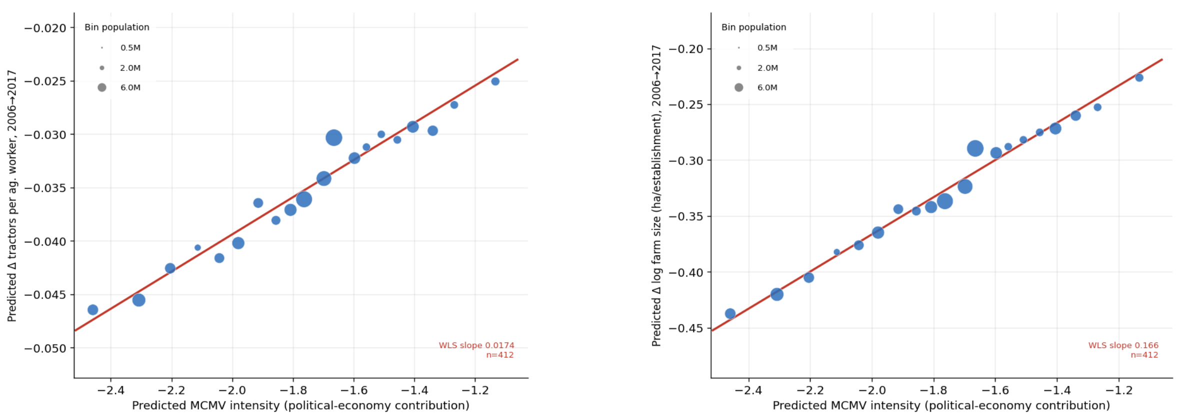


(a) Δ log agricultural labour

(b) Δ log agricultural wage

Notes: This figure shows binned scatters of MCMV's *predicted* housing program intensity, at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and *predicted* decadal changes in agricultural organization outcomes, measured by the 1997, 2006, and 2017 wages of the Agricultural census, using the same predictors. Both axes use the same set of political economy predictors and control for the same set of baseline socio-economic controls, defined in Section 2.4.2. Binned means are population-weighted; red line is weighted-least squares fit over $x \leq$99th percentiles; markers are proportional to bin population.

Figure 7: Predicted mechanization and farm-scale changes vs. predicted MCMV intensity.



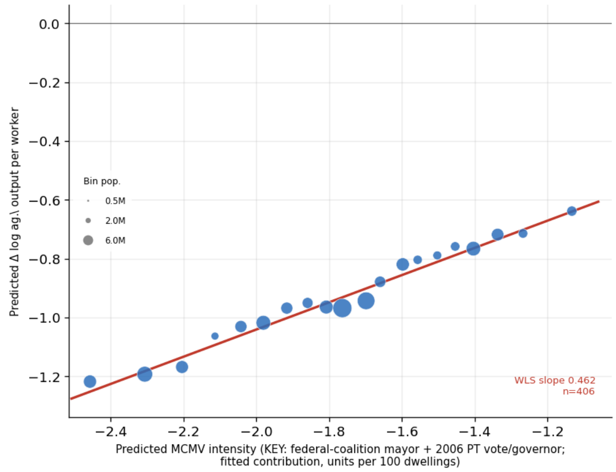
(a) Δ tractors per worker

(b) Δ log farm size

Notes: See footnotes to Figure 6.

SICONV corroborates the pattern. Discretionary federal infrastructure transfers (SICONV) display the same agricultural gradients as MCMV, and the effect is concentrated in urban-housing and sanitation spending—not rural infrastructure transfers, which show no relationship with agricultural reorganization (Appendix C.2, Figures C1–C2). Since the two predicted intensities share the same political-economy predictors and are therefore correlated (Appendix Figure C3), the single-instrument columns in Tables 5–6 are the more interpretable comparison; the SICONV results serve primarily to confirm that the agricultural reorganization is a feature of construction-driven urban labor demand broadly, not an artifact of MCMV's specific institutional rules.

Figure 8: Predicted agricultural output per worker vs. predicted MCMV intensity.



Notes: See notes to Figure 6. Output per worker is measured as total production value in the microregion, valeud at (constant 2017 R$) per the PAM dataset divided by total agricultural workers in the microregion, per the AgCensus data.

2.4.4 Heterogeneity by Baseline Characteristics

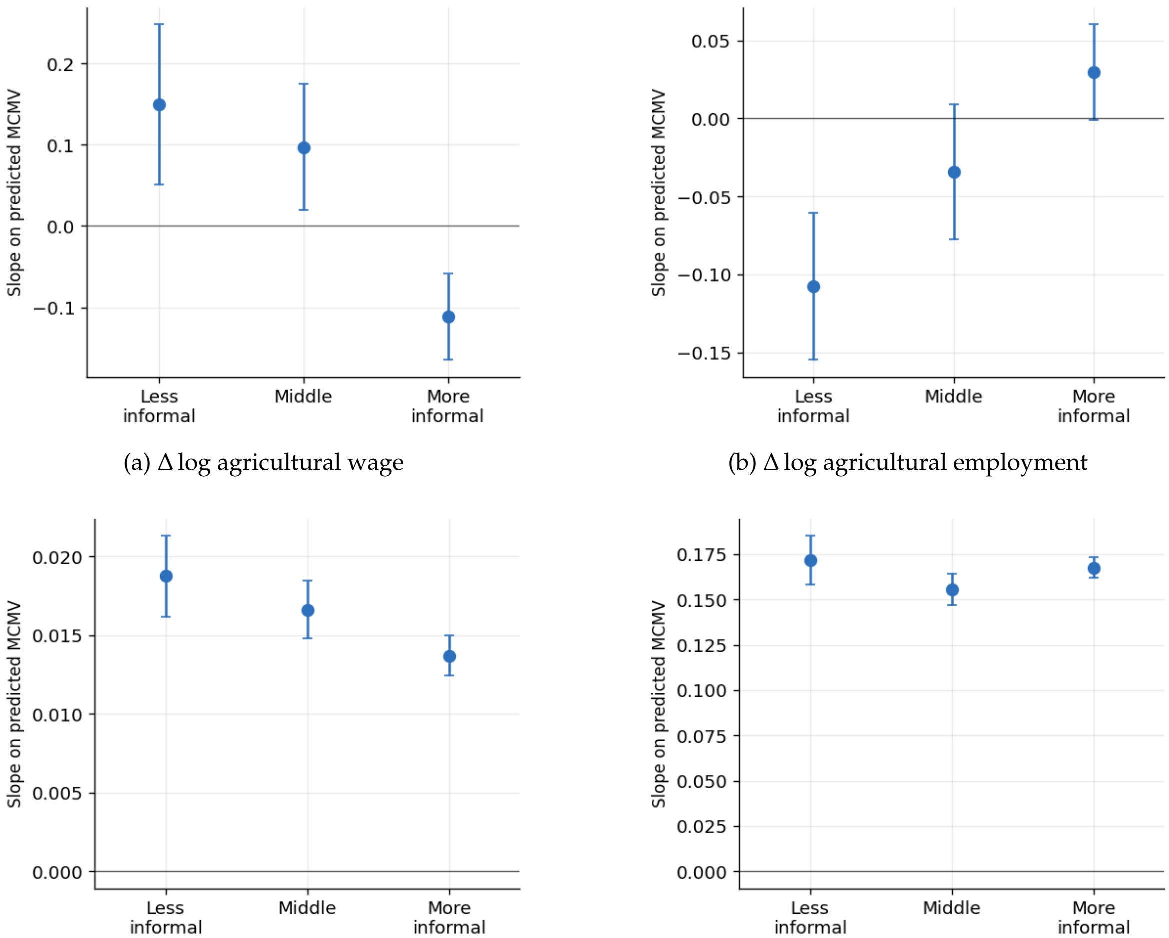
The two heterogeneity dimensions we study test the transmission mechanism rather than just the sign of the effect. Both ask: given that a construction shock arrives, what determines how much of it reaches the agricultural labor market and is converted into farm reorganization?

Informality as a labor buffer. Lewis (1954)'s model of structural transformation assigns a central role to the informal sector as a reservoir of labor. When formal non-agricultural demand rises, workers flow from informal into formal employment first; only once this reservoir is depleted do agricultural workers face genuinely better outside options and does the relative price of agricultural labor begin to rise. In microregions where the non-farm economy is predominantly informal, the labor demand shock is absorbed by this informal buffer rather than by pulling workers directly from agriculture. The relative cost of agricultural labor therefore rises less, and the incentive to reorganize agricultural production is correspondingly weaker. The mechanism is ultimately about relative prices: informality keeps the relative cost of labor in agriculture from changing.

Figure 9 plots the predicted-on-predicted slopes by tercile of the 2000 non-agricultural informality rate (the share of prime-age non-farm workers in informal jobs). The gradient is consistent with the Lewis account at every step of the chain. In the most-informal tercile, the agricultural wage and mechanization slopes are essentially flat; in the least-informal tercile, they are steep and precisely estimated.

The continuous interaction in Table D1 confirms the pattern: a one-standard-deviation increase in non-agricultural informality reduces the agricultural wage slope by 0.11, the agricultural employment slope by 0.06, and the tractors-per-worker slope by 0.002—all statistically significant. The evidence is thus consistent with the reallocation margin operating through the formal labor market: where informal slack mutes

Figure 9: Predicted outcome changes vs. predicted MCMV intensity, by baseline non-ag informality



(a) Δ log agricultural wage

(b) Δ log agricultural employment

(c) Δ tractors per worker

(d) Δ log farm size

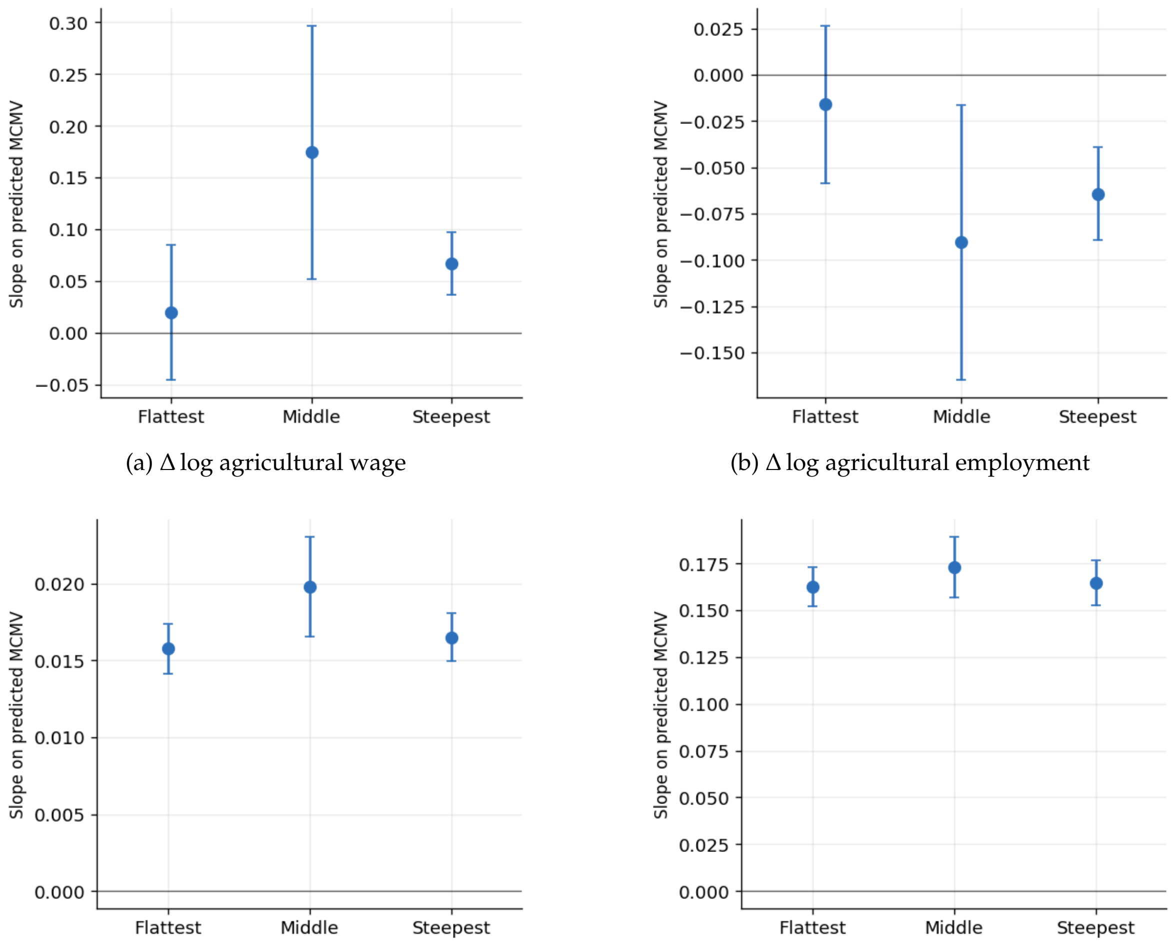
Notes: Predicted-on-predicted fits within terciles of the 2000 *non-agricultural* informal-employment share. This is the share of prime-age (18–64) workers employed in non-agricultural industries who hold non-formal jobs, from the 2000 Census microdata aggregated to the microregion. SE clustered at the microregion. Tercile cutoffs are printed at the bottom of each panel; continuous interactions in Table D1.

the wage signal, farms see little change in the relative price of labor and adjust correspondingly less.

One outcome is notably robust to informality: farm consolidation. The farm-size gradient is stable across all informality terciles, and the continuous interaction is small and insignificant (Table D1). This robustness points to an important distinction between the consolidation and mechanization channels. Mechanization requires a substantial increase in the relative cost of labor to be profitable. By contrast, any wage increase that pushes operating costs above the variable profit of the marginal farm is sufficient to drive it out of business. Even in high-informality areas where the wage increase is attenuated, a modest rise in labor costs may be enough to eliminate the least-productive operations, with their land absorbed by surviving farms. Consolidation is therefore the most robust channel of agricultural reorganization—the one that operates even when the wage and mechanization channels are muted.

Terrain and the extensive margin of mechanization. Terrain constrains the feasibility and returns to tractor adoption: flat land enables machinery, steep slopes inhibit it. The natural prior is monotone—flatter terrain should support larger responses to a wage-driven mechanization incentive. Figure 10 shows the opposite: the response is largest at *intermediate* terrain steepness for all four outcomes—wages, employment, tractors, and farm size.

Figure 10: Predicted outcome changes vs. predicted MCMV intensity, by terrain-steepness tercile.



(a) Δ log agricultural wage

(b) Δ log agricultural employment

(c) Δ tractors per worker

(d) Δ log farm size

Notes: Microregion level; each panel plots the predicted-on-predicted fits within terciles of mean terrain slope (Flattest / Middle / Steepest). No macro-region fixed effects in the second stage (they would absorb the topographic variation of interest); SE clustered at the microregion. Tercile cutoffs in the variable's units are printed at the bottom of each panel.

The non-monotonic pattern is consistent with an extensive-margin story about mechanization thresholds. In the flattest microregions, farms were already predominantly mechanized at the 2006 baseline; having crossed the adoption threshold, they have limited scope for further capital deepening in response to the MCMV shock. In the steepest microregions, the physical constraint on mechanization means that the relative price increase does not translate into tractor adoption regardless of magnitude. In intermediate-terrain microregions, mechanization is feasible but not yet prevalent at baseline: farms have the topographic capacity to mechanize and the unrealized potential to do so, and a construction-driven increase in the cost of labor provides the decisive price signal to cross the threshold. This is an extensive-margin effect—what matters is the change in whether to mechanize, not how much.

The fact that the non-monotonicity extends to agricultural wages, employment, and farm size, not just tractors, suggests that intermediate-terrain areas contained at the 2006 baseline a disproportionate share of "on-the-margin" farms: operations large enough to survive historically but not efficient enough to absorb a wage increase. The construction shock pushed these farms below their zero-profit threshold, inducing exit and consolidation while raising average wages and productivity among surviving operations. Consistent with this interpretation, the linear terrain interaction in Table D1 is small and mostly insignificant: a monotone slope is not the right specification for an effect driven by threshold-crossing rather than by continuous variation in topography.

Together, informality and terrain locate two distinct bottlenecks in the transmission chain. Informality governs whether the wage signal reaches agriculture at all—a Lewis-style buffer that weakens the relative price change at the source. Terrain governs how agriculture responds conditional on receiving the signal—an extensive-margin filter that determines which farms have the technical capacity to substitute capital for labor and which are simply pushed out of business. Both bottlenecks are quantitatively important and economically interpretable, and their joint patterns provide support for the mechanism at the core of the model in the Section 3.

3 A Model of Firms, Farms and Input Intensification

We develop a multi-sector general-equilibrium model to interpret the reduced-form evidence and identify the mechanism through which non-agricultural labor-demand shocks drive structural transformation in agriculture. The model departs from canonical structural change frameworks in two ways. First, farm mechanization is endogenous: farms choose how many production tasks to perform with capital versus labor in response to relative factor prices. Second, this choice is subject to an irreversibility: expanding the set of tasks that can be performed with capital requires a one-time, fixed technology-adaptation cost that is not recovered upon reversal. Together, these features generate the persistent reorganization of agricultural production documented in Section 2.4—mechanization, farm exit, land consolidation, and higher output per worker—even in response to a temporary construction shock.

3.1 Preferences and Technology

Environment. Time is discrete and infinite. A measure-one continuum of identical households supplies labor inelastically. The aggregate labor endowment is $\bar{N}$, which is allocated across productive activities and overhead costs of firm operation and entry. Markets are complete.

The economy has three sectors. *Non-agriculture* consists of heterogeneous firms that produce with labor alone. *Agriculture* consists of heterogeneous farms that produce with labor, capital, and a fixed factor (land). An *informal sector* (services) employs a representative firm with a fixed technology. Non-agricultural and informal output are perfect substitutes in consumption; agricultural output is a distinct good with a subsistence requirement. The total supply of land is fixed at $\bar{L}$, usable only in agriculture. Capital K_t accumulates endogenously through household saving.

Preferences. Households have Stone-Geary preferences over agricultural good C_{at}, non-agricultural good C_{nt}, and informal/services good C_{st}:

$$W \equiv \max_{C_{nt},C_{at},C_{st},b_{t+1}} \sum_{t=0}^{\infty} \nu^t \Big[\theta_n \ln\big(C_{nt} - \bar{c}_n\big) + \theta_a \ln\big(C_{at} - \bar{c}_a\big) + \theta_s \ln\big(C_{st} - \bar{c}_s\big)\Big], \tag{4}$$

where $\nu \in (0,1)$ is the discount factor, $\bar{c}_n, \bar{c}_a, \bar{c}_s \geq 0$ are subsistence levels, and $\theta_n + \theta_a + \theta_s = 1$. As income rises, expenditure shares shift toward non-agricultural goods (the Engel-curve channel of structural transformation). We treat the planner's problem directly; shadow prices λ_o correspond to sector-o goods and μ_N to the shadow cost of labor.

Households face the following budget constraint:

$$P_{nt}C_{nt} + P_{at}C_{at} + P_{st}C_{st} + b_{t+1} \leq W_t\bar{N}_t + R_{bt}b_t + R_tL_t + \pi_{at} + \pi_{nt} + \pi_{st}, \tag{5}$$

$$b_0 \geq 0, \tag{6}$$

where π_n, π_s, π_a are the profits associated to the firms and agricultural farms, L is the land endowment; and b_{t+1} are savings with return R_{bt}. For simplicity, we assume no land depreciation.

Technologies: Non-agricultural Sector. Non-agricultural firms are heterogeneous in productivity Z_{ni}, drawn upon entry from distribution G_n. Firm i's output is

$$Y_{ni} = Z_{ni} N_{ni}^{\alpha}, \quad \alpha \in (0,1). \tag{7}$$

Entry requires paying a fixed cost F_n in units of labor. Active firms pay a per-period operating cost f_n in labor units and exit at exogenous rate δ_{Mn}. Free entry drives the expected value of entry to zero. The marginal active firm has threshold productivity $\bar{Z}_n$; all firms with $Z_{ni} \geq \bar{Z}_n$ operate. The measure of active firms M_{nt} evolves as

$$M_{n,t+1} = m_{nt}\big(1 - G(\bar{Z}_{n,t+1})\big) + M_{nt}(1 - \delta_{Mn}), \tag{8}$$

where m_{nt} is the mass of entrants.

An informal sector employs a representative firm with technology $Y_{st} = Z_s N_{st}^{\alpha}$ and no entry or operating costs.

3.2 The Agricultural Sector

Farms are heterogeneous in productivity Z_{aj}, drawn from distribution G_a upon entry. Each farm uses labor N_{aj}, capital K_{aj}, and land L_{aj}, and exits at exogenous rate δ_{Ma} or endogenously upon hitting the zero-profit threshold.

3.2.1 Production Technology and Task Aggregation

Production at farm j is organized as a unit continuum of tasks $b \in [0,1]$, aggregated with a Cobb-Douglas technology:

$$Y_{aj} = Z_{aj} \exp\!\left(\gamma \int_0^1 \ln x_j(b)\,\mathrm{d}b\right) L_{aj}^{\gamma_L}, \qquad \gamma + \gamma_L < 1, \tag{9}$$

where $x_j(b)$ denotes the output of task b at farm j.

Tasks differ in the relative efficiency of labor versus capital. Let $a_n(b)$ be the labor-augmenting efficiency of task b and $a_k(b)$ the capital-augmenting efficiency.

Assumption 1 (Comparative advantage) *$a_n(b)/a_k(b)$ is continuous and strictly decreasing in b. Labor has comparative advantage in low-b tasks; capital has comparative advantage in high-b tasks.*

Whether a task can be performed with capital depends on the farm's *mechanizable set*. Let $\tilde{\beta}_j \in [0,1]$ be the lower boundary of farm j's mechanizable set: tasks $b \in [\tilde{\beta}_j, 1]$ can be performed with either capital or labor, while tasks $b \in [0, \tilde{\beta}_j)$ can only use labor. Thus:

$$x_j(b) = \begin{cases} a_n(b)\, n_j(b) + a_k(b)\, k_j(b) & b \in [\tilde{\beta}_j,\, 1], \\ a_n(b)\, n_j(b) & b \in [0,\, \tilde{\beta}_j), \end{cases} \tag{10}$$

where $n_j(b)$ and $k_j(b)$ are labor and capital inputs to task b.

3.2.2 Reduced-form Production Function

Because $x_j(b)$ enters in logs and all tasks receive equal expenditure share (Appendix E.1), factor inputs are equalized within each class: $n_j(b) = N_{aj}/\beta$ for all labor tasks and $k_j(b) = K_{aj}/(1-\beta)$ for all capital tasks, where $\beta \in [\tilde{\beta}_j, 1]$ is the threshold activity performed by labor. Substituting into equation (9) yields the reduced-form production function

$$Y_{aj} = Z_{aj}\, A(\beta)\, \big(K_{aj}^{1-\beta}\, N_{aj}^{\beta}\big)^{\gamma}\, L_{aj}^{\gamma_L}, \tag{11}$$

where $A(\beta) \equiv \exp\Big(\gamma\big[\int_0^{\beta} \ln a_n(b)\,\mathrm{d}b + \int_{\beta}^{1} \ln a_k(b)\,\mathrm{d}b\big]\Big)\, \beta^{-\gamma\beta}(1-\beta)^{-\gamma(1-\beta)}$ is the comparative-advantage efficiency term. Here $1-\beta$ is the capital share and β is the labor share within the composite input $(\,\cdot\,)^{\gamma}$.

Mechanization with Fixed Costs. The mechanizable set $\tilde{\beta}_j$ is a state variable for each farm. At entry, $\tilde{\beta}_j = 1$: the mechanizable set is empty (measure zero), and all tasks must use labor. Expanding the mechanizable set—reducing $\tilde{\beta}_j$ below its current level—requires paying a fixed adaptation cost I in units of agricultural output (e.g., redrawing field plots, purchasing specialized equipment). This cost is not incurred when the mechanizable set contracts.

Optimal mechanization: no cost benchmark. In the absence of fixed costs ($I = 0$), the farm sets $\tilde{\beta}_j = \beta$, so the mechanizable and actual mechanization thresholds coincide. The optimal threshold satisfies the indifference condition

$$\frac{a_n(\beta^*)}{a_k(\beta^*)} = \frac{\mu_N}{\lambda_k}, \tag{12}$$

where μ_N is the shadow cost of labor and λ_k is the shadow rental rate of capital. Because $a_n(b)/a_k(b)$ is strictly decreasing (Assumption 1), equation (12) has a unique solution $\beta^*(\mu_N/\lambda_k)$ that is decreasing in the wage-rental ratio: higher wages induce more mechanization (lower β^*, fewer labor tasks).

Optimal mechanization: with fixed costs. When $I > 0$, the farm does not mechanize task b unless the capital cost savings net of the adaptation cost are positive. For the marginal task at the mechanizable threshold $\tilde{\beta}$, this break-even condition is

$$\big[p_x(\tilde{\beta})\, a_k(\tilde{\beta}) - \lambda_k\big]\, \bar{k} = \lambda_a I, \tag{13}$$

where $\bar{k} = K_{aj}/(1-\tilde{\beta})$ is the capital input per mechanizable task in a symmetric equilibrium. Using the optimality condition for labor in adjacent tasks, $p_x(\tilde{\beta})\, a_n(\tilde{\beta}) = \mu_N$, and rearranging, equation (13) becomes

$$\frac{a_n(\tilde{\beta})}{a_k(\tilde{\beta})} = \frac{\mu_N}{\lambda_k + \lambda_a I/\bar{k}}, \tag{14}$$

where $\bar{k} = K_{aj}/(1-\tilde{\beta})$. Since $\lambda_k + \lambda_a I/\bar{k} > \lambda_k$, comparing equations (12) and (14) and using the strict monotonicity of a_n/a_k yields the following result.

Proposition 1 (Fixed costs reduce mechanization) *For any $I > 0$, the mechanization threshold with fixed costs satisfies $\tilde{\beta} > \beta^*$: fewer tasks are performed with capital than in the no cost economy. Moreover, (i) the gap $\tilde{\beta} - \beta^*$ is increasing in I: implicitly differentiating (14), $\partial F/\partial I = \mu_N \lambda_a/[\bar{k}(\lambda_k + \lambda_a I/\bar{k})^2] > 0$ while $\partial F/\partial \tilde{\beta} = (d/d\tilde{\beta})[a_n/a_k] < 0$ by Assumption 1, so the implicit-function theorem gives $\partial \tilde{\beta}/\partial I = -(\partial F/\partial I)/(\partial F/\partial \tilde{\beta}) > 0$. (ii) The gap is decreasing in K_{aj}: larger K_{aj} raises $\bar{k} = K_{aj}/(1-\tilde{\beta})$, reducing $\lambda_a I/\bar{k}$ and hence the RHS of (14); by the same implicit-function argument, $\partial \tilde{\beta}/\partial K_{aj} < 0$, so larger farms mechanize more. Higher agricultural goods prices λ_a also widen the gap, since the fixed cost is more expensive in value terms*

Within the mechanizable set, the farm optimally uses capital for tasks $b \in [\beta, 1]$ and labor for tasks $b \in [\tilde{\beta}, \beta)$, where β satisfies the standard indifference condition $a_n(\beta)/a_k(\beta) = \mu_N/\lambda_k$. Because $\beta \geq \tilde{\beta} > \beta^*$, not all mechanizable tasks are necessarily exercised with capital: the farm may keep some mechanizable tasks in labor even when $\tilde{\beta}$ is below the frictionless optimum.

3.2.3 Irreversibility and Persistent Mechanization

The key departure from no fixed cost benchmark is that the fixed cost I is paid only when the mechanizable set expands—that is, only when $\tilde{\beta}$ decreases. Once a set of tasks has been made mechanizable, that technological adaptation persists at zero additional cost regardless of subsequent factor prices. Formally, the mechanizable threshold evolves as

$$\tilde{\beta}_{j,t+1} \leq \tilde{\beta}_{j,t}, \tag{15}$$

with equality if the farm does not expand its mechanizable set in period t (no cost incurred), and strict inequality if it does (cost I incurred).

This irreversibility generates hysteresis in farm technology: a temporary increase in the relative cost of labor leads to a permanently larger mechanizable set, producing persistent mechanization even after the wage shock dissipates.

Proposition 2 (Temporary shocks generate permanent mechanization) *Consider a farm* at the mechanization constraint *in the initial equilibrium: $\beta_0 = \tilde{\beta}_0$ (the farm is constrained, since the frictionless optimum $\beta_0^* < \tilde{\beta}_0$ is unattainable without paying I). A temporary rise in μ_N/λ_k induces the farm to pay I and expand the mechanizable set to $\tilde{\beta}' = \beta^*(\mu_N^1) < \beta_0^* < \tilde{\beta}_0$. After the shock dissipates and factor prices return to (μ_N^0, λ_k^0), two* *permanent changes remain. First, the mechanizable set is permanently larger: $\tilde{\beta}$ stays at $\tilde{\beta}' < \tilde{\beta}_0$ (irreversibility constraint (15)). Second, actual mechanization permanently improves: the farm can now reach the frictionless optimum β_0^* freely (since $\tilde{\beta}' < \beta_0^*$), so β falls permanently from $\tilde{\beta}_0$ to $\beta_0^* < \tilde{\beta}_0$.*

Proof See Appendix E.2.

The intuition is straightforward. When wages rise temporarily, the farm finds it profitable to pay I and lower $\tilde{\beta}$ from $\tilde{\beta}_0$ to some $\tilde{\beta}' < \tilde{\beta}_0$, enabling more tasks to use capital. When wages normalize, the frictionless optimum β^* rises back toward $\tilde{\beta}_0$. Under the irreversibility constraint, $\tilde{\beta}'$ is a permanent lower bound on the mechanizable threshold: the farm can set $\beta \geq \tilde{\beta}'$ freely, and for any $\beta \in (\tilde{\beta}', \tilde{\beta}_0)$ the farm now benefits from capital without paying additional adaptation costs. The result is more mechanization than would have prevailed absent the shock—even after the shock is gone.

Remark Proposition 2 implies that even a *temporary* construction boom—such as the MCMV program studied in Section 2.4—can generate a permanent shift in the technology mix of the agricultural sector. The temporary rise in the relative cost of labor induces farms near the mechanization margin to pay the fixed adaptation cost and reconfigure production; that reconfiguration is not reversed when construction activity subsides.

3.2.4 Farm Optimization and Exit

Each farm maximizes the present discounted value of profits. In a symmetric equilibrium, farm j's per-period profit (after paying operating cost f_a and possibly the mechanization fixed cost I) is

$$\pi_{aj}(Z_{aj}) = \lambda_a Y_{aj} - \mu_N N_{aj} - \lambda_k K_{aj} - \mu_L L_{aj} - f_a \mu_N - \mathbf{1}[\tilde{\beta} \text{ falls}] \cdot \lambda_a I. \tag{16}$$

At the zero-profit threshold $\bar{Z}_a$, profit equals zero, which pins down the cutoff productivity. Using the first-order conditions (see Appendix F) and the fact that capital-labor and land-labor ratios are equalized across farms with the same $\tilde{\beta}$, the zero-profit condition can be written as

$$(1-\gamma-\gamma_L)\, \lambda_a K_a^{(1-\beta)\gamma} N_a^{\beta\gamma} L_a^{\gamma_L} Z_a \left(\frac{\bar{Z}_a}{Z_a}\right)^{\frac{1}{1-\gamma-\gamma_L}} = f_a \mu_N, \tag{17}$$

where Z_a is the average productivity of the marginal firm and $\bar{Z}_a$ is the threshold. A rise in μ_N (higher cost of labor) raises the left-hand side only if it does not fully pass through to λ_a and λ_k—the standard result that higher wages erode farm profitability, raising $\bar{Z}_a$ and inducing exit of the least productive farms. The measure of active farms M_{at} evolves analogously to equation (8).

Proposition 3 (Farm consolidation) *Suppose μ_N rises while the change in λ_a (agricultural goods price) is less than proportionate—which holds when agricultural goods are traded internationally at world prices, or more generally whenever the non-agricultural shock does not fully pass through into agricultural goods prices. Then the* *zero-profit condition (17) implies that $\bar{Z}_a$ rises (the marginal farm's operating cost exceeds its revenue), the measure of active farms M_{at} falls, and average farm size $\bar{L}/M_{at}$ increases. This consolidation channel operates independently of whether terrain allows mechanization: even farms that cannot mechanize face tighter margins when labor costs rise relative to the agricultural goods price.*

Proposition 3 is the model's explanation for why farm-size consolidation is the most robust dimension of agricultural reorganization in the data: unlike mechanization (which requires topographic feasibility) or wage responses (which require a tight formal labor market), the zero-profit mechanism requires only that some farms operate at the margin.

3.3 Equilibrium

3.3.1 Planner's Problem

We characterize the equilibrium through the planner's problem since the economy is efficient. The planner maximizes household welfare (4) subject to feasibility constraints:

$$C_{at} = M_{at}\,\bar{Y}_{at}, \tag{18}$$

$$C_{nt} + X_t = M_{nt}\,\bar{Y}_{nt}, \tag{19}$$

$$C_{st} = Y_{st}, \tag{20}$$

$$K_{t+1} = X_t + (1-\delta_k)\,K_t, \tag{21}$$

$$M_{i,t+1} = m_{it}\big(1 - G(\bar{Z}_{i,t+1})\big) + M_{it}(1-\delta_{Mi}), \quad i \in \{a, n\}, \tag{22}$$

$$N_t^p + N_t^f = \bar{N}, \tag{23}$$

where X_t is investment, N_t^p is productive labor (employed in production), $N_t^f \equiv F_a m_{at} + F_n m_{nt} + M_{at}\int_{\bar{Z}_a} f_a\, g_a\, dZ + M_{nt}\int_{\bar{Z}_n} f_n\, g_n\, dZ$ is overhead labor (entry and operating costs), $\bar{Y}_{it}$ is the output of the average active firm in sector i, and δ_{Mi} is the exogenous exit rate.

The planner's problem attaches shadow prices λ_a, λ_n, λ_s to goods constraints (18)–(20); μ_X to the capital accumulation constraint (21); ψ_a, ψ_n to the gricultural and non-agricultural firm-measure evolution constraints (22); and μ_N to the labor market constraint (23). The full set of first-order conditions is in Appendix F.

3.3.2 Market Clearing and Equilibrium Definition

Stationary equilibrium A stationary equilibrium is a collection of allocations $\{C_a, C_n, C_s, X, N_{aj}(Z), K_{aj}(Z), L_{aj}(Z), N_{ni}(Z), N_s, m_a, m_n, \beta, \tilde{\beta}\}$, prices $\{\lambda_a, \lambda_n, \lambda_s, \mu_N, \lambda_k, \mu_L\}$, thresholds $\{\bar{Z}_a, \bar{Z}_n\}$, and measures $\{M_a, M_n\}$ such that:

1. Households maximize (4) taking prices as given.

2. Farms maximize profits taking $\tilde{\beta}$ as a state variable and subject to the irreversibility constraint (15).

3. Non-agricultural firms maximize profits.

4. Goods, labor, capital, and land markets clear:

$$M_a \bar{Y}_a = C_a, \tag{24}$$

$$M_n \bar{Y}_n = C_n + X, \tag{25}$$

$$M_n \int_{\bar{Z}_n} N_{ni}(Z)\, g_n(Z)\, dZ + N_s + M_a \int_{\bar{Z}_a} N_{aj}(Z)\, g_a(Z)\, dZ + N^f = \bar{N}, \tag{26}$$

$$M_a \int_{\bar{Z}_a} K_{aj}(Z)\, g_a(Z)\, dZ = K, \tag{27}$$

$$M_a \int_{\bar{Z}_a} L_{aj}(Z)\, g_a(Z)\, dZ = \bar{L}. \tag{28}$$

5. Free entry is satisfied in both sectors: expected value of entry equals entry cost.

3.3.3 Equilibrium Characterization

We characterize the stationary equilibrium. Let $s \equiv X/(M_n \bar{Y}_n)$ denote the investment rate, and let $\bar{Z}_n$ and $\bar{Z}_a$ denote the productivity thresholds in non-agriculture and agriculture.

Note on derivations. Equations (29)–(31) below follow from combining the first-order conditions in Appendix F with the Stone-Geary expenditure conditions, goods-market clearing, and the Pareto structure of G_n and G_a. Specifically, the Pareto property that average above-threshold output is a constant multiple of threshold output (e.g. $\bar{Y}_n \propto \bar{Y}_n(\bar{Z}_n)$ with ratio $\alpha_n(1-\alpha)/(\alpha_n(1-\alpha)-1)$) plays a key role.

Non-agricultural employment. From the planner's optimality conditions and free entry,

$$M_n^* N_n^* = \frac{N^p}{1 + (1-s) + \dfrac{\gamma\beta}{\alpha} \cdot \dfrac{\theta_n(1-s)}{\theta_a}}, \tag{29}$$

where N_n^* is employment of the average non-agricultural firm. A reduction in non-agricultural entry costs (an increase in M_n) raises total non-agricultural employment directly and reduces agricultural employment through the labor-market clearing condition.

Agricultural employment relative to non-agriculture.

$$\frac{M_a^* N_a^*}{M_n^* N_n^*} = \frac{\gamma\beta}{\alpha} \cdot \frac{\theta_a}{\theta_n(1-s)}. \tag{30}$$

A rise in μ_N induces mechanization (lower β), directly reducing the labor share in agriculture and shifting the relative employment ratio.

Free entry in non-agriculture.

$$M_n^* = \frac{\theta_n(1-\alpha)}{\hat{\delta}_{Mn}} \cdot \frac{1-G(\bar{Z}_n)}{\nu\, F_n\, \mu_N}, \tag{31}$$

where $\hat{\delta}_{Mn} = 1 - \nu(1-\delta_{Mn})$. An exogenous decline in the non-agricultural entry cost F_n raises M_n^*, increases μ_N through labor-market pressure, and triggers the reorganization of agricultural production characterized by Propositions 1–3.

Capital accumulation. The Euler equation for capital satisfies

$$\mu_{X,t} = \nu[\mu_{X,t+1}(1-\delta_k) + \mu_{K,t+1}], \tag{32}$$

where $\mu_{K,t+1} = \lambda_{a,t+1}\, \partial\bar{Y}_{a,t+1}/\partial K$ is the shadow marginal product of capital in agriculture. In the steady state, the capital-output ratio in agriculture satisfies

$$1 = \nu\left[(1-\delta_k) + \frac{\lambda_a(1-\beta)\gamma\bar{Y}_a}{\lambda_n K_a}\right]. \tag{33}$$

Relative factor prices and the labor-productivity relationship. Using the first-order conditions for labor in each sector and the static allocation across firms, the ratio of average labor productivities between agriculture and the informal sector satisfies

$$\frac{\bar{Y}_a/N_a}{\bar{Y}_s/N_s} = \frac{\alpha\,\lambda_s}{\gamma\beta\,\lambda_a}, \tag{34}$$

which links sectoral labor productivities to relative goods prices and the degree of mechanization.

Cross-firm relative employment. Within each sector, the ratio of employment across two farms with productivities Z_{aj} and $Z_{aj'}$ satisfies

$$\frac{Z_{aj}}{Z_{aj'}} = \left(\frac{N_{aj}}{N_{aj'}}\right)^{1-\gamma-\gamma_L}, \tag{35}$$

so more productive farms are larger and—since land is equally divided and capital ratios equalized—operate more efficiently per unit of labor.

4 Quantitative Analysis

In this section, we present the quantification of our model. We use the same data sources from our empirical analysis to estimates key moments that discipline the steady state equilibrium allocation, and targets the characteristics of the Brazil economy between 1986 and 1995. We then calibrate a change in entry costs in the non-agriculture sector that mimic the observed raise in real wages, consistent with our reduced form findings. Given this calibrated shock, we assess the magnitude of employment reallocation, labor productivity, mechanization and land reallocation predicted by the model and compare it with the observed patters from section 2.3.2. Finally, we evaluate the impact of the same shock under counterfactual scenarios where farms lack the ability to become more capital intensive as relative prices change.

Current implementation: two simplifying assumptions. First, we assume the economy is at the *no cost mechanization optimum*: farms always operate at $\beta = \beta^*$ (the no-cost benchmark from (12)), so the fixed adaptation cost I is irrelevant for the current quantitative exercise. The irreversibility constraint and the mechanization-gap dynamics of Section 3.2.3 are therefore switched off. Second, the current exercise solves for a *stationary equilibrium* (steady state), comparing model-implied long-run moments to their data counterparts. Transition dynamics—the intended final quantitative exercise—are left for a subsequent version; the steady-state solution here serves as the starting point and terminal condition for that computation.

4.1 Model Parameterization

We take a hybrid approach to quantifying 21 parameters that fully describe the equilibrium of the model, including factor endowments, technology distributions, entry and operating costs as well as preferences. We take seven standard parameters from the literature, estimate four parameters from our data before simulating the model, and set three parameters exogenously, including a normalization for productivity in the informal agriculture. Then, we jointly calibrate the remainder six parameters to match moments in the data. Since we work with the stationary equilibrium, we parameterize the profile of comparative advantage that underlies the mechanization decision so that $\mathcal{A}(\beta) = 1$ in the steady state.[6]

Table 7: Parameterization

Parameter	Description	Value	Source/Moment
α	Share of labor in non-agriculture	0.82	Ákos Valentinyi and Herrendorf (2008)
γ_l	Share of land in agriculture	0.18	Ákos Valentinyi and Herrendorf (2008)
γ	Share of equipment and labor composite in agriculture	0.71	Ákos Valentinyi and Herrendorf (2008)
ι	Comparative advantage in agr.	2	Acemoglu and Zilibotti (2001)
ν	Discount factor	0.96	Annual interest rate of 4%
δ_k	Depreciation rate of equipment in agriculture	0.0294	Caunedo and Keller (2020)
δ_n	Exit rate in non-agriculture	13%	RAIS data, Ulyssea (2018)
δ_a	Exit rate in agriculture	11%	RAIS data
θ_a	Share of agriculture in consumption	0.2	National Accounts
$\theta_s = \theta_n$	Share of non-agriculture in consumption	0.4	National Accounts
f_n	Operating cost non-agriculture	$0.5w$	Ulyssea (2018)
f_a	Operating cost agriculture	$0.1w$	Exogenous, below f_n
F_n	Entry cost in non-agriculture	0.0016	Exogenous
Z_s	Productivity in informal non-agriculture	1	Normalization

Notes: Benchmark model parameterization.

Parameters directly calibrated from the literature. For the production technologies we use the labor share $\alpha = 0.82$ in non-agricultural sectors and for the agriculture, we set the land share to be equal $\gamma_l = 0.18$

[6] This model calibration is preliminary and in progress.

with the composite share of labor and capital to be 0.7. These parameters are taken from Ákos Valentinyi and Herrendorf (2008) and correspond to factor shares estimated for the US using detailed national accounts information. The depreciation rate of equipment in agriculture is set as in Caunedo and Keller (2020). The exit rate in non-agriculture is taken from Ulyssea (2018) and estimated using RAIS data. Similarly to Ulyssea (2018) we set the operating cost in non-agriculture exogenously at 50% of the wage, and we set exogenously the operating cost in agriculture to 10% of the wage. Finally, we parameterize the profile of comparative advantage as in Acemoglu and Zilibotti (2001): $\frac{a_n(b)}{a_k(b)} \equiv \phi\left(\frac{b}{1-b}\right)^{\iota-1}$, for $\iota = 2$, and ϕ chosen so that the equilibrium labor share correspond to the calibrate one, given capital-output ratios.[7]

Parameters inferred directly from the data. We set the discount factor $\nu = 0.96$ to reflect a real interest rate of 4%. To estimate the share of consumption for each sector we look at the relative shares in the data, so $\theta_a = 0.2$. For model tractability we calibrate $\theta_n = \theta_s = (1 - \theta_a)/2$, so that $\theta_a + \theta_s + \theta_n = 1$. Finally, we normalize productivity in informal non-agricultural sector to 1, so that $Z_s = 1$. Table 7 summarizes parameters calibrated directly from the literature, as well as estimated from the data.

Parameters estimated by matching moments The rest of the parameters are internally calibrated by jointly matching moments that inform the remaining parameters values. The parameters we need to calibrate are the following. Tail parameter of Pareto distribution in each sector, α_n and α_a, as well as entry cost in agriculture. To obtain tail parameters of distribution in agriculture and non-agriculture we target share of large farms and firms observed in the data, respectively. We rely on the data from RAIS for non-agricultural firms and target share of firms with employment larger that than average employment in this sector. For agriculture sector, we target the share of farms with land size above 100 ha, since individual level data are not available and we have to rely on bins provided by Ag Census. To obtain the value of entry cost, we ratio of labor productivity in agriculture to the one in non-agriculture for the entire economy.

In addition, we estimate total land and labor endowments for the whole economy, L and $\bar{N}$, as well as the share of labor in the value added composite, β. Total land endowment is estimated using information on average land size of farms across municipalities from Ag Census. And to estimate total labor endowment we target the share of labor in agriculture for the entire economy using data from the World Bank. Finally, to calibrate share of labor within composite for agriculture, we target capital to value added value in agriculture computed using data World Input Output Database for Brazil. Parameter values and respective data and model moments are depicted in Table 8.

[7] We first calibrate the expenditure shares in an economy without mechanization, and then invert the model to infer the underlying shape of the profile of comparative advantage.

Table 8: Parameters Calibrated Jointly by Matching Moments

Moments			Model Parameters		
Description	Target	Model	Var.	Value	Description
Share of non-agr. firms with above average empl.	12.3%	12.5%	α_n	6.8	Productivity distr. non-agr., tail parameter
Share of farms larger than 100 ha	10.7%	9.6%	α_a	9	Productivity distr. agr., tail parameter
Average land per farm (*ha*)	72.7	73.7	L	180	Total land endowment
Share of labor in agriculture	15.8%	15.4%	$\bar{N}$	70	Total labor endowment
Labor productivity in agr. to non-agr.	0.25	0.28	F_a	0.6	Entry cost in agriculture
Capital to output ratio in agriculture	2.17	2.14	β	0.85	Share of labor within composite

Notes: Benchmark model parameterization, moments calibrated jointly.

4.2 Model Mechanisms

After calibrating the model to the steady state that corresponds to the Brazilian economy in 90s, we further use observed raise in wages from our empirical analysis to estimate the effect of large entry shock in the model. To do so, we lower the entry cost in non-agricultural sector such that it leads to the change in wage as observed in the data and compare new steady state to the baseline model steady state.

4.2.1 Large Firm Entry Shock in the Model

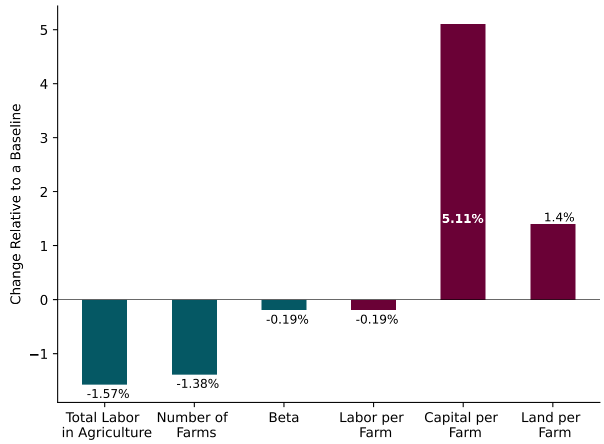
Similarly to the empirical analysis, we use the model to study the shock that arises from the entry of large non-agricultural firms and its consequent effects on labor allocation and agricultural organization. Specifically, the model examines how a reduction in the entry cost for non-agricultural firms F_n triggers structural adjustments in the economy, particularly within the agricultural sector. We reduce F_n by 19.5 percent, so that this change triggers wage increase of 1.4% percent corresponding to the one observed in the data over the decade following a large entry shock.

A reduction in F_n incentivizes firm entry in the non-agricultural sector, leading to an expansion in labor demand outside of agriculture. As employment opportunities increase, wages rise, making labor in agriculture more expensive relative to other inputs. This shift induces a decline in agricultural employment, as farmers decrease the demand for labor as a result of higher labor cost. Consequently, the total number of farms decreases, with the contraction predominantly concentrated among smaller farms, which are not able to compete in a labor market with rising wages.

As small farms exit, the average landholding size per farm increases, leading to a process of land consolidation. Simultaneously, rising labor costs induce farms to adopt more capital-intensive production technologies. This shift results in increased mechanization, as farms substitute labor with capital, leading to a higher number of tractors per hectare and per farm. The model thus captures an endogenous response to labor market shocks, wherein farms adjust their production structure in response to changing factor prices.

Quantitatively, the model predicts that labor reallocation away from agriculture follows from the higher

Figure 11: The Effect of Entry Shock on Organization of Agriculture



Notes: Figure depicts the effect of 19.5 percent reduction in entry cost in non-agricultural sector (consistent with wage increase of 1.4 percent) on organization of agriculture relative to the baseline steady state in the model with endogenous mechanization.

relative labor cost for farmers. The exit of smaller farms, coupled with fixed land supply, drives a consolidation process that alters the overall structure of agricultural production. In response to labor shortages, surviving farms mechanize and increase capital intensification, further reinforcing the transformation of the agricultural sector. Figure 11 depicts main outcomes in agriculture predicted by the model as a result of large entry shock.

These model-driven insights align with the empirical findings, where the entry of large non-agricultural firms leads to a measurable shift in employment patterns, farm size distribution, and capital use in agriculture. Through comparative statics and counterfactual analysis, the model provides a framework for understanding the role of external labor market shocks in driving structural transformation.

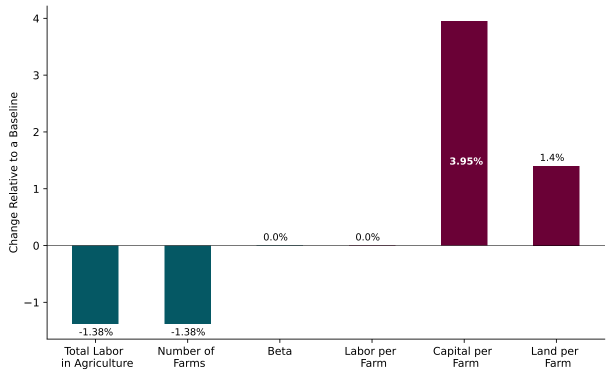
4.2.2 The Role of Mechanization for Agricultural Responses

To understand the role of mechanization in shaping agricultural responses to non-agricultural shocks, we contrast two theoretical environments from our model: one in which mechanization is endogenous described in previous Section, and another where mechanization channel is shut down and the level of capital intensity remains unchanged.

A large non-agricultural firm entry—modeled as a reduction in the fixed entry cost induces a positive labor demand shock in the non-agricultural sector. This raises wages economy-wide and triggers adjustments in the agricultural sector via both labor reallocation and production reorganization. However, the magnitude and nature of these responses differ significantly depending on whether agricultural mechanization is allowed to adjust.

When mechanization is not permitted to respond, the agricultural sector lacks the flexibility to offset

Figure 12: The Effect of Entry Shock on Organization of Agriculture (No Mechanization)



Notes: Figure depicts the effect of reduction in entry cost in non-agricultural sector consistent with wage increase of 1.4 percent on organization of agriculture relative to the baseline steady state in the model with fixed labor share in agriculture.

labor shortages with capital deepening. First, labor reallocation from agriculture is more limited. Differently from a standard model of structural change (where all sectors are equally intensive in labor), labor in agriculture still declines because of differential labor intensity across sectors, as in Acemoglu and Guerrieri (2008). While farm exit and higher capital usage per farm still occurs as a result of change in relative prices of inputs, labor per farm essentially remains the same. Since no capital intensification is allowed, labor reallocation from agriculture is driven solely by farms exit, and hence decline by the same magnitude. The main model predictions for agricultural sector are depicted in Figure 12.

Without mechanization, the agricultural sector's capacity to absorb shocks is constrained. These comparisons illustrate that endogenous mechanization amplifies the pull effects of non-agricultural growth on structural transformation. In its absence of capital adoption, adjustments are slower and more reliant on farm exits.

4.3 Counterfactuals *Work-in-progress*

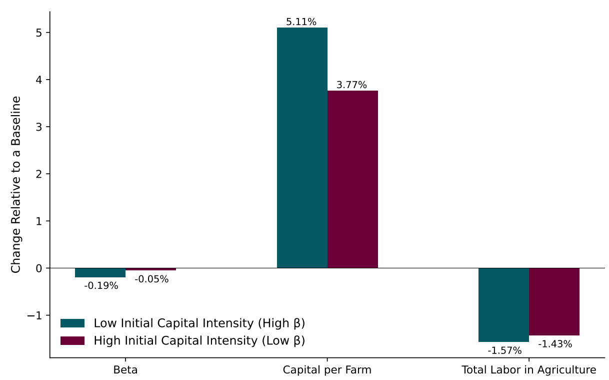
Building on our baseline counterfactual, we plan to extend our quantitative analysis by incorporating initial heterogeneity in agricultural organization across regions. Brazil exhibits substantial variation in the structure of agricultural production—particularly in the functioning of land and capital markets, prevalence of smallholder versus large-scale farms, and initial levels of mechanization. For example, regions such as the South and Southeast are characterized by relatively more consolidated farms and higher baseline levels of capital intensity, while the North and Northeast tend to rely more heavily on labor-intensive agriculture and exhibit limited mechanization. These differences imply that identical labor market shocks may generate heterogeneous responses across space, depending on the flexibility of production systems and factor substitutability.

To explore this, we plan to recalibrate the model to reflect regional characteristics, varying key parameters such as land endowments, capital intensity, and the distribution of farm productivity. We will then simulate the same large firm entry shock across these alternative calibrations to study how the magnitude and channels of structural adjustment vary given this heterogeneity. This exercise will allow us to assess whether endogenous mechanization plays a more prominent role in areas with stronger capital markets, and conversely, whether reallocation relies more on the extensive margin (e.g., farm exit and land consolidation) in settings where mechanization is more constrained. These additional counterfactuals will help validate the external validity of our framework and offer deeper insights into the regional dynamics of structural transformation, with important implications for agricultural development strategies in different settings.

4.3.1 Large Firm Entry Shock in Environment with Heterogeneous Initial Mechanization

To illustrate how heterogeneity in the initial level of mechanization affects the economic response to a labor demand shock, we repeat the exercise from Section 4.2.1 for a model in which agriculture is initially more mechanized–that is, where β is lower. More specifically, we reduce the fixed cost in non-agriculture to the level that generates the same wage increase as in our baseline exercise, or as observed in the data (1.4 percent).[8] The main model outcomes for this higher-mechanization case are shown in Figure 13.

Figure 13: The Effect of Entry Shock on Agriculture Based on Initial Level of Mechanization in Agriculture



Notes: Figure depicts the effect of reduction in entry cost in non-agricultural sector consistent with wage increase of 1.4 percent on organization of agriculture relative to the baseline steady state in the model with higher and lower initial level of mechanization, i.e. lower or higher β. The low level of mechanization depicts our baseline calibration.

The model predicts that an economy with more mechanized agriculture exhibits more muted effects from the entry shock than one where agriculture is more labor-intensive. Because agriculture already relies heavily on capital, the scope for further capital intensification–and, consequently, for reorganization of production–is more limited. As a result, we observe smaller changes in labor reallocation and in capital per farm.

[8]Alternatively, we can apply the same shock in terms of the magnitude of the reduction in entry costs in non-agriculture. Qualitatively, both approaches lead to similar outcomes.

By contrast, in a model with a fixed labor share in agriculture, the quantitative agricultural responses to the same wage increase are identical in economies with more and less mechanized agriculture. In that framework, the primary channel of adjustment operates through changes in relative prices, so a given wage increase produces the same proportional changes in agricultural outcomes regardless of the initial mechanization level. Thus, allowing for endogenous mechanization is essential for capturing how identical shocks translate into systematically different outcomes across heterogeneous economies.

5 Conclusion

This paper asks what happens to the organization of agricultural production when non-agricultural labor demand expands. The answer, documented across two quasi-experimental designs in Brazil, is that farms reorganize substantially: wages rise, employment falls, capital intensity increases, farms consolidate, and average productivity grows. These changes are not a mechanical byproduct of workers leaving agriculture—they reflect active adjustments in how the farms that remain choose to produce.

An important finding is that the sector composition of non-agricultural growth matters. Construction entries drive the agricultural response; manufacturing entries of comparable size do not. The reason is straightforward: in contemporary Brazil, agriculture and construction overlap heavily in the low end of the skill distribution, while manufacturing has shifted toward more educated workers. This suggests that, at the current stage of Brazilian development, the mechanism linking non-agricultural growth to agricultural transformation operates primarily through construction and related low-skill services—not through manufacturing.

The effects are persistent. Agricultural wages for low-education workers at construction epicenters continue to grow a decade after entry, and the agricultural reorganization documented over 2006–2017 has no counterpart in the pre-program period. The model provides one explanation for this persistence beyond the wage channel: once a farm invests in expanding its mechanizable task set, that investment does not unwind when factor prices normalize. Fixed adaptation costs, combined with irreversibility, imply that even episodic construction booms can leave a lasting imprint on agricultural technology.

Two heterogeneity results sharpen the mechanism. In microregions with more informal non-agricultural employment, the agricultural response is attenuated, consistent with the informal sector absorbing the demand shock before it reaches agricultural wages. In microregions at intermediate terrain steepness, the mechanization and consolidation responses are largest, consistent with an extensive-margin story in which the construction shock provides the decisive incentive to cross the mechanization threshold for farms that had the capacity but not yet the price signal. Farm consolidation is the one channel that operates robustly across both dimensions—even modest wage increases are enough to push marginal farms below the zero-profit threshold.

Taken together, the evidence points to a view of agricultural transformation in which the non-agricultural economy does not merely absorb surplus labor from agriculture, but actively reshapes it. Whether this dynamic is specific to Brazil or reflects a broader pattern in middle-income economies where manufacturing has become skill-intensive while construction has not is a question we leave for future work. The current analysis provides the empirical foundation and the theoretical framework to investigate it.

References

Acemoglu, D. and V. Guerrieri (2008): "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 116, 467–498.

Acemoglu, D. and P. Restrepo (2019): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33, 3–30.

Acemoglu, D. and F. Zilibotti (2001): "Productivity Differences*," *The Quarterly Journal of Economics*, 116, 563–606.

Adamopoulos, T. and D. Restuccia (2014): "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, 104, 1667–97.

Alvarez-Cuadrado, F. and M. Poschke (2011): "Structural Change Out of Agriculture: Labor Push versus Labor Pull," *American Economic Journal: Macroeconomics*, 3, 127–58.

Amodio, F., P. Medina, and M. Morlacco (2025): "Labor Market Power, Self-Employment, and Development," *American Economic Review*, 115, 3014–3057.

Asher, S. and P. Novosad (2020): "Rural Roads and Local Economic Development," *American Economic Review*, 110, 797–823.

Boppart, T., P. Kiernan, P. Krusell, and H. Malmberg (2023): "The Macroeconomics of Intensive Agriculture," CEPR Discussion Papers 18039, C.E.P.R. Discussion Papers.

Bustos, P., B. Caprettini, and J. Ponticelli (2016): "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, 106, 1320–65.

Bustos, P., J. Castro-Vincenzi, J. Monras, and J. Ponticelli (2019): "Industrialization without Innovation," Working Paper 25871, National Bureau of Economic Research.

Caunedo, J. and N. Kala (2021): "Mechanizing Agriculture," Working Paper 29061, National Bureau of Economic Research.

Caunedo, J. and E. Keller (2020): "Capital Obsolescence and Agricultural Productivity*," *The Quarterly Journal of Economics*, 136, 505–561.

De Chaisemartin, C. and X. d'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American economic review*, 110, 2964–2996.

Dix-Carneiro, R. and B. K. Kovak (2017): "Trade liberalization and regional dynamics," *American Economic Review*, 107, 2908–2946.

Engbom, N. and C. Moser (2022): "Earnings Inequality and the Minimum Wage: Evidence from Brazil," *American Economic Review*, 112, 3803–47.

Felix, M. (2026): "Trade, Labor Market Concentration, and Wages," Working Paper 35018, National Bureau of Economic Research.

Foster, A. D. and M. R. Rosenzweig (2010): "Microeconomics of Technology Adoption," *Annual Review of Economics*, 2, 395–424.

——— (2022): "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size," *Journal of Political Economy*, 130, 636–680.

Gollin, D., S. Parente, and R. Rogerson (2007): "The food problem and the evolution of international income levels," *Journal of Monetary Economics*, 54, 1230–1255.

Harris, J. R. and M. P. Todaro (1970): "Migration, Unemployment and Development: A Two-Sector Analysis," *The American Economic Review*, 60, 126–142.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2014): "Growth and Structural Transformation," in *Handbook of Economic Growth*, Elsevier, vol. 2, chap. 06, 855–941, 1 ed.

Hopenhayn, H. A. (1992): "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Lewis, W. A. (1954): "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 22, 139–191.

Madhok, R., F. Noack, A. M. Mobarak, and O. Deschenes (2025): "Internal Migration and the Spatial Reorganization of Agriculture," Working Paper 33854, National Bureau of Economic Research.

Manuelli, R. and A. Seshadri (2014): "Frictionless Technology Diffusion: The Case of Tractors," *American Economic Review*, 104, 1368–91.

Méndez, E. and D. Van Patten (2022): "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company," *Econometrica*, 90, 2685–2721.

Peters, M., Y. Zhang, and F. Zilibotti (2026): "Skipping the Factory: Service-Led Growth and Structural Transformation in the Developing World," NBER Working Papers 34692, National Bureau of Economic Research, Inc.

Rodrik, D. (2016): "Premature deindustrialization," *Journal of Economic Growth*, 21, 1–33.

Schultz, T. (1953): "The Economic Development of Our Western Interior," *American Journal of Agricultural Economics*, 35, 707–713.

SHARMA, G. (2025): "Collusion Among Employers in India," Working paper, revise and resubmit, American Economic Review.

SIMONI JUNIOR, S. AND E. DIAS (2025): "A dimensão Partidária-Eleitoral do Programa Minha Casa Minha Vida," *Revista DADOS*, 68.

STORESLETTEN, K., B. ZHAO, AND F. ZILIBOTTI (2026): "Business cycle during structural change: Arthur Lewis' theory from a neoclassical perspective," *The Economic Journal*, ueag008.

ULYSSEA, G. (2018): "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, 108, 2015–47.

ZEIRA, J. (1998): "Workers, Machines, and Economic Growth*," *The Quarterly Journal of Economics*, 113, 1091–1117.

ÁKOS VALENTINYI AND B. HERRENDORF (2008): "Measuring factor income shares at the sectoral level," *Review of Economic Dynamics*, 11, 820–835.

A Appendix Tables and Figures

Table A1: Characteristics of the analyzed large entries, by entry sector

Characteristic	Construction / real estate	Manufacturing	Other
Entries and size			
Number of large entries	57	15	9
Total employees at entry	28,639	5,653	3,409
Mean employees at entry	502	377	379
Median employees at entry	334	282	260
Entry size, % of microregion pop. (mean)	0.22%	0.16%	0.20%
Region (number of entries)			
North	9	0	1
Northeast	16	7	2
Southeast	18	6	5
South	4	2	1
Central-West	10	0	0
Firm / establishment			
Entirely new firm	49%	20%	56%
New establishment, existing firm	51%	80%	44%
Origin of workers at entry			
From another formal (RAIS) job	90%	96%	92%
From outside formal employment	10%	4%	8%
Education of workers at entry			
Primary	18%	13%	11%
Secondary	75%	82%	83%
Tertiary	7%	5%	7%
Demographics of workers at entry			
Male	81%	65%	72%
Aged $\leq$24	29%	26%	10%
Aged 25–39	49%	55%	49%
Aged 40–64	22%	19%	40%
Aged 65+	0%	0%	1%
White (of race-recorded)	35%	68%	66%
Black (of race-recorded)	5%	4%	4%
Brown/*parda* (of race-recorded)	60%	27%	29%
Asian or Indigenous (of race-rec.)	0%	0%	0%

Notes: One row per epicenter municipality (the earliest qualifying entry, $N = 81$): a new non-agricultural establishment with $\geq$200 employees at entry and $\geq$0.1% of microregion population, in 1998–2017, excluding public administration, agriculture, and agricultural-demand industries (food/beverage, textile, rubber/tobacco), and excluding microregions with a large entry before 1998 — the same selection as the difference-in-differences sample. "Construction / real estate" and "Manufacturing" (IBGE transformation industries) are the two largest sectors; *Other* collects the remaining qualifying entries — here transport and communication, hospitality, mining, and wholesale trade. Region is the five IBGE macro-regions (from the municipality code). Worker shares (origin, education, demographics) are of the entrant establishment's workforce at entry, pooled across entries; "from another formal (RAIS) job" vs. "from outside formal employment" splits the entry headcount by whether the worker held a formal job in the prior year. Race shares are of workers with race recorded (race is missing for older entries).

Table A2: Cross-sectional incidence of large entries on MCMV and SICONV-housing intensity

	Constr./RE entry (0/1)		Other entry (0/1)	
	(1)	(2)	(3)	(4)
MCMV units / 100 dwellings	0.00519***	0.00518***	0.000986	0.000976
	(0.0019)	(0.0019)	(0.00095)	(0.00095)
SICONV housing R$ per capita		0.00012		0.000122
		(9.3e-05)		(9e-05)
Log population (2010)	0.0602***	0.0607***	0.0212***	0.0217***
	(0.0087)	(0.0089)	(0.0055)	(0.0056)
Region FE	Yes	Yes	Yes	Yes
N (MCAs)	1278	1278	1278	1278

Notes: Linear probability models at the MCA level. Dependent variable: indicator for having an epicenter large entry ($\geq$200 employees) in construction/real estate (cols. 1–2) or in another sector (cols. 3–4). MCMV is units contracted (2009–2014) per 100 dwellings; SICONV housing is committed *habitacao* repasse (2008–2017) in R$ per capita; both winsorized at the 99th percentile. All columns include log 2010 population and region fixed effects — indicators for Brazil's five regions (North, Northeast, Southeast, South, Central-West), each MCA assigned the modal region of its municipalities. Heteroskedasticity-robust (HC1) standard errors in parentheses. */**/***: $p < 0.10/0.05/0.01$.

Table A3: Predictors of MCMV intensity (units per 100 dwellings), by contracting window.

	All contracts 2009–2014		*Mayor elections Oct 2008* Contracts 2009–2012		*President/Gov. election Oct 2010* Contracts 2011–2014	
	(1) all	(2) Faixa 1	(3) all	(4) Faixa 1	(5) all	(6) Faixa 1
Program design predictors						
Municipal housing deficit, 2010 (%)	0.0535	0.0219	0.0104	-0.00217	0.0333	0.0046
	(0.033)	(0.03)	(0.025)	(0.024)	(0.021)	(0.015)
State housing deficit, 2010 (%)	0.0186	0.109**	0.0196	0.0705*	0.00809	0.0848**
	(0.069)	(0.053)	(0.043)	(0.036)	(0.049)	(0.036)
Metropolitan region, 2010	2.46***	0.443	1.32***	0.239	2.11***	0.316
	(0.81)	(0.39)	(0.45)	(0.25)	(0.73)	(0.31)
SICONV housing, 2008–09 (R$/res.)	0.00258	0.00317	0.00295	0.00325	0.0018	0.00239
	(0.0036)	(0.0033)	(0.0031)	(0.003)	(0.0025)	(0.0023)
Political economy predictors (elections)						
Mayor in federal coalition (2009)	0.716***	0.422***	0.55***	0.383***	0.589***	0.33***
	(0.25)	(0.14)	(0.15)	(0.1)	(0.2)	(0.096)
Core PT presidential vote, 1994–2006	-0.0122	0.0124	-0.011	0.000585	-0.0132	0.0123
	(0.023)	(0.015)	(0.015)	(0.011)	(0.019)	(0.011)
Swing PT presidential vote, 1994–2006	-0.0237	0.0125	-0.00127	0.0251	-0.0335	-0.00916
	(0.033)	(0.026)	(0.025)	(0.019)	(0.024)	(0.02)
PT state governor, 2006	0.39	0.533	-0.0239	0.104	0.526	0.632*
	(0.48)	(0.39)	(0.29)	(0.24)	(0.4)	(0.32)
Baseline socio-economic correlates (2000)						
Urban population share	1.21	-1.61**	0.903	-0.883*	1.17	-1.18**
	(1.1)	(0.68)	(0.68)	(0.49)	(0.85)	(0.5)
<50,000 inhabitants	-3.62***	-2.94***	-2.55***	-1.68***	-1.88***	-1.63***
	(0.54)	(0.4)	(0.31)	(0.25)	(0.44)	(0.26)
Agricultural employment share	0.894	-0.225	0.665	-0.251	0.165	-0.654
	(1.3)	(0.89)	(0.84)	(0.67)	(1.1)	(0.71)
Manufacturing employment share	0.746	-4.14***	-0.157	-3.25***	0.688	-2.77***
	(2.4)	(1.1)	(1.4)	(0.76)	(2.1)	(0.91)
Formal employment share	1.85	-1.07	1.71	0.257	-0.081	-2.53***
	(1.6)	(1.2)	(1.1)	(0.98)	(1.2)	(0.74)
N (MCAs)	3658	3658	3658	3658	3658	3658
R^2	0.091	0.103	0.096	0.067	0.060	0.102

Notes: MCA (Minimum Comparable Area) level; OLS without geographic fixed effects; standard errors clustered at the microregion (mmc) level in parentheses (413 clusters). Dependent variable: MCMV units contracted per 100 dwellings (units and `Domicilios_total` summed across the MCA's municipalities), over three contracting windows (odd columns: all units; even columns: Faixa 1). *Program-design predictors*: the municipal and state 2010 FJP relative housing deficit (the state deficit, the legal allocation level, is total state deficit over total state dwellings); a metropolitan-region indicator (2010); and SICONV housing (habitação) transfers per resident signed in 2008–09 (a prior-capacity proxy). *Political-economy predictors*: a mayor in the federal PT-led coalition (2009); the core and swing PT presidential vote (mean/SD, 1994–2006); and a PT state-governor indicator (2006). *Baseline socio-economic correlates* (2000): urban population share, a <50,000-inhabitants indicator, and the agricultural, manufacturing, and formal employment shares (a construction share is unavailable: the 2000-Census municipal panel has only agriculture/manufacturing/services). All regressors are 2010-population-weighted MCA means of their municipal values. */**/***: $p < 0.10/0.05/0.01$.

Table A4: MCMV allocation regressions, reproducing Simoni Jr. & Dias (2025), Tables 1–3.

	(1) MCMV all units	(2) MCMV Faixa 1 Faixa 1
Panel A. Mayor in federal coalition — contracts 2009–2012		
Mayor in federal coalition (2009)	0.195	0.346***
	(0.12)	(0.13)
Housing deficit, 2010 (% of dwellings)	-0.0102	-0.0163
	(0.013)	(0.014)
<50,000 inhabitants, 2010	-2.71***	-3.07***
	(0.4)	(0.43)
Metropolitan region, 2010	0.586	-0.0929
	(0.49)	(0.35)
Urbanisation rate, 2010 (%)	0.00765	-0.023
	(0.0091)	(0.015)
Controls (log GDP p.c. 2010, HDI 2010), Region FE	Yes	Yes
N / R^2	5564 / 0.121	5564 / 0.098
Panel B. Core/swing PT vote (1994–2006) — contracts 2009–2010		
Core PT presidential vote	0.0048	0.00351
	(0.0098)	(0.011)
Swing PT presidential vote	-0.0059	-0.00318
	(0.0059)	(0.0069)
Housing deficit, 2010 (% of dwellings)	0.00789	0.00737
	(0.0071)	(0.0079)
<50,000 inhabitants, 2010	-1.65***	-2.04***
	(0.26)	(0.29)
Metropolitan region, 2010	0.247	0.166
	(0.21)	(0.23)
Urbanisation rate, 2010 (%)	0.00574	-9.21e-05
	(0.005)	(0.0065)
Controls (log GDP p.c. 2010, HDI 2010), Region FE	Yes	Yes
N / R^2	5565 / 0.110	5565 / 0.089
Panel C. Core/swing PT vote (1994–2010) — contracts 2011–2014		
Core PT presidential vote	0.0192	0.0524**
	(0.025)	(0.025)
Swing PT presidential vote	-0.0036	-0.00855
	(0.039)	(0.035)
Housing deficit, 2010 (% of dwellings)	0.0118	0.00736
	(0.017)	(0.018)
<50,000 inhabitants, 2010	-2.31***	-3.35***
	(0.46)	(0.64)
Metropolitan region, 2010	0.964	-0.303
	(0.65)	(0.34)
Urbanisation rate, 2010 (%)	0.0042	-0.0479**
	(0.011)	(0.019)
Controls (log GDP p.c. 2010, HDI 2010), Region FE	Yes	Yes
N / R^2	5565 / 0.094	5565 / 0.122

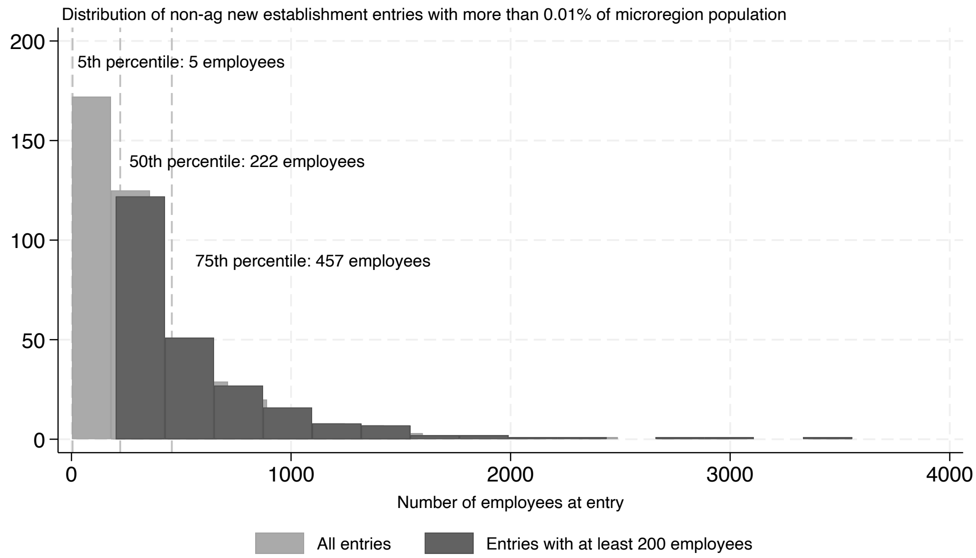
Notes. Municipality level; OLS reproducing the MCMV-distribution model of Simoni Jr. & Dias (2025) (DADOS 68(1), Tables 1–3). Dependent variable: MCMV units contracted over the indicated window per 100 dwellings — all-units columns normalised by total dwellings, Faixa 1 columns by Faixa-1 (low-income) dwellings. Each panel is a separate model: Panel A uses the mayor-in-federal-coalition indicator (the mayor's party in the federal PT-led coalition); Panels B and C use the core and swing PT presidential vote (mean and standard deviation of the PT vote over 1994–2006 for the 2009–2010 contracts and over 1994–2010 for the 2011–2014 contracts). Program-design variables: relative housing deficit (2010), a <50,000-inhabitants indicator (2010), a metropolitan-region indicator (2010), and the urbanisation rate (2010). Controls (reported as Yes): log GDP per capita (2010) and the municipal HDI (2010); all panels include region fixed effects. Standard errors clustered by state (UF) in parentheses. $^*/^{**}/^{***}$: $p < 0.10/0.05/0.01$.

Table A5: Predicted MCMV intensity and predicted outcome changes: collated binscatter slopes.

Outcome (predicted change)	Slope on predicted MCMV	R^2	N
Δ log agric. wage (Ag. Census), 2006–17	0.0875*	0.09	402
	(0.046)		
Δ log agric. labour (Ag. Census), 2006–17	-0.0505**	0.09	402
	(0.025)		
Δ tractors per worker (Ag. Census), 2006–17	0.0174***	0.81	402
	(0.0013)		
Δ log farm size (Ag. Census), 2006–17	0.166***	0.91	402
	(0.0062)		
Δ log construction employment (RAIS), 2009–17	0.303***	0.27	393
	(0.065)		
Δ log agric. employment (RAIS), 2009–17	-0.109***	0.62	397
	(0.0092)		
Δ log primary-educ. wage (RAIS), 2009–17	0.0197	0.04	402
	(0.014)		
Δ log secondary-educ. wage (RAIS), 2009–17	0.00626	0.01	402
	(0.0093)		

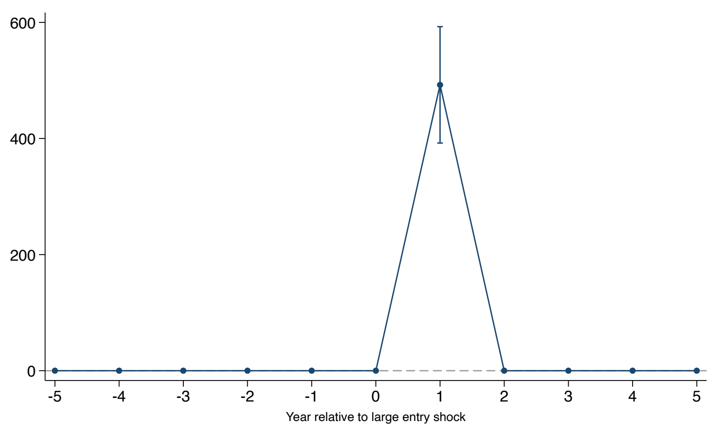
Notes: Microregion (mmc) level; one row per outcome — the regression analogue of the binscatters above. The dependent variable is the *predicted* outcome change and the regressor is *predicted* MCMV intensity: each is the fitted contribution of the political-economy predictors (a mayor in the federal coalition 2009, the 2006 PT presidential core/swing vote, and a PT state-governor indicator 2006), from an OLS of the target on those predictors plus the 2000 socio-economic controls (urban population share, a $<$50,000-inhabitants indicator, the 2005 metropolitan-region indicator, and the agricultural, manufacturing, and formal employment shares) and macro-region (5) fixed effects, retaining only the political-economy contribution. The RAIS construction-employment row is formal construction employment (Constr_eqwor) summed to the microregion and then logged, so the microregion-level aggregate is positive and no observations are dropped. The slope reproduces the population-weighted (WLS) binscatter slope over the 1st–99th percentile of predicted MCMV intensity. Because both sides are deterministic functions of the same predictors, this is a *descriptive* predicted-on-predicted relationship, not a causal estimate. Estimated by WLS weighted by 2000 population; standard errors clustered by state (UF) in parentheses. $^{*}/^{**}/^{***}$: $p < 0.10/0.05/0.01$.

Figure A1: Size distribution of qualifying non-agricultural establishment entries



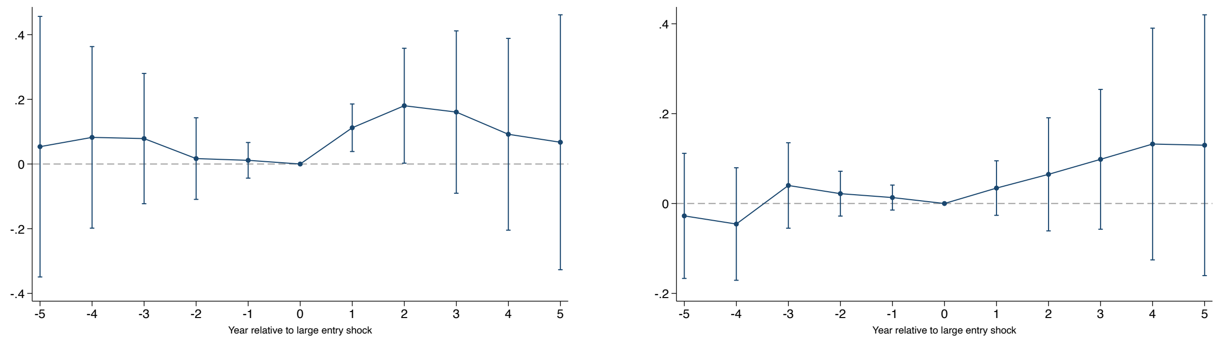
Notes: Distribution of employees at entry for new non-agricultural establishment entries exceeding 0.1% of microregion population. Light bars: all such entries; dark bars: entries with at least 200 employees (the qualifying cutoff). Dashed lines mark the 5th (5), 50th (222), and 75th (457) percentiles.

Figure A2: Entry-shock size: epicenter vs. never-entered



Notes: Cumulative event-study path of formal employees at entry, epicenter vs. never-entered municipalities (a mechanical check on the size of the shock).

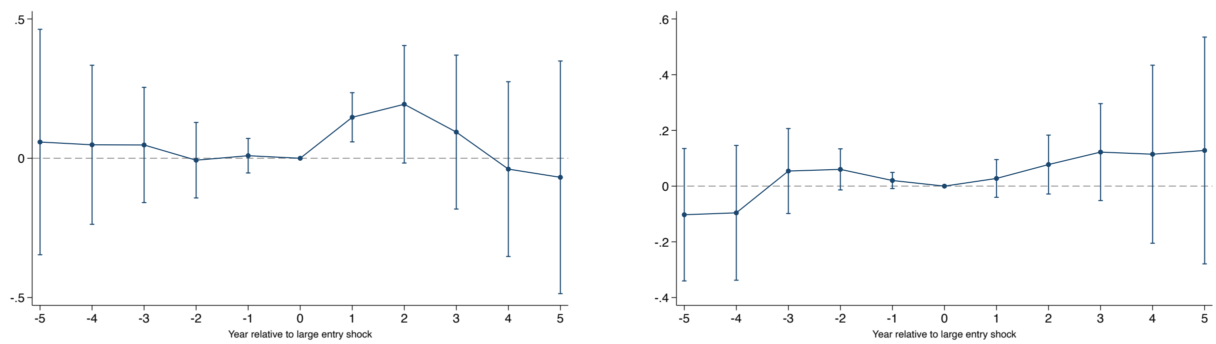
Figure A3: Log total formal employment — by entry sector



(a) Construction / real estate (b) Manufacturing

Notes: Epicenter vs. never-entered, cumulative `did_multiplegt_dyn` event study (`placebo(5) effects(10)`, `trends_lin`, `only_never_switchers`); SEs clustered at the MMC level; event years -5 to $+10$; the horizontal line marks zero, with the treated (epicenter) side restricted to a single entry sector (construction/real estate or manufacturing, the IBGE transformation industries) against the same never-entered controls.

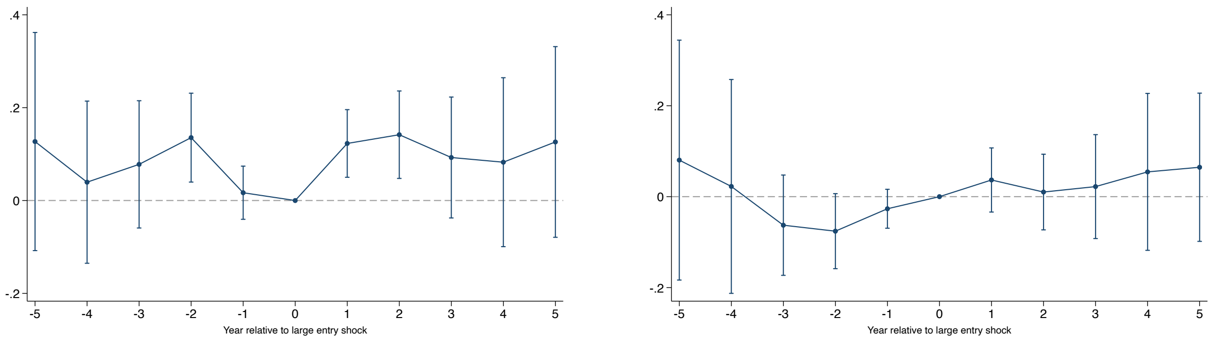
Figure A4: Log formal employment, primary education — by entry sector



(a) Construction / real estate (b) Manufacturing

Notes: Epicenter vs. never-entered, cumulative `did_multiplegt_dyn` event study (`placebo(5) effects(10)`, `trends_lin`, `only_never_switchers`); SEs clustered at the MMC level; event years -5 to $+10$; the horizontal line marks zero, with the treated (epicenter) side restricted to a single entry sector against the same never-entered controls.

Figure A5: Log real monthly wage, primary education — by entry sector

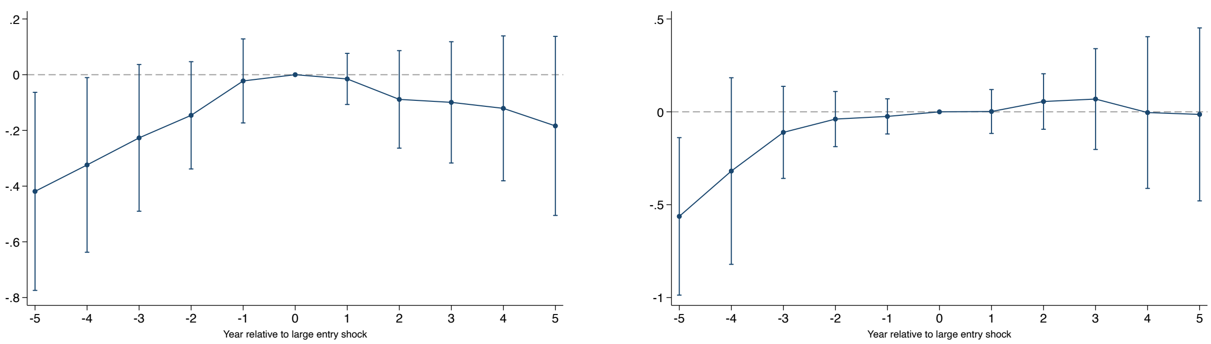


(a) Construction / real estate

(b) Manufacturing

Notes: Epicenter vs. never-entered, cumulative `did_multiplegt_dyn` event study (`placebo(5) effects(10)`, `trends_lin`, `only_never_switchers`); SEs clustered at the MMC level; event years -5 to $+10$; the horizontal line marks zero, with the treated (epicenter) side restricted to a single entry sector (construction/real estate or manufacturing, the IBGE transformation industries) against the same never-entered controls.

Figure A6: Log formal agricultural employment — by entry sector

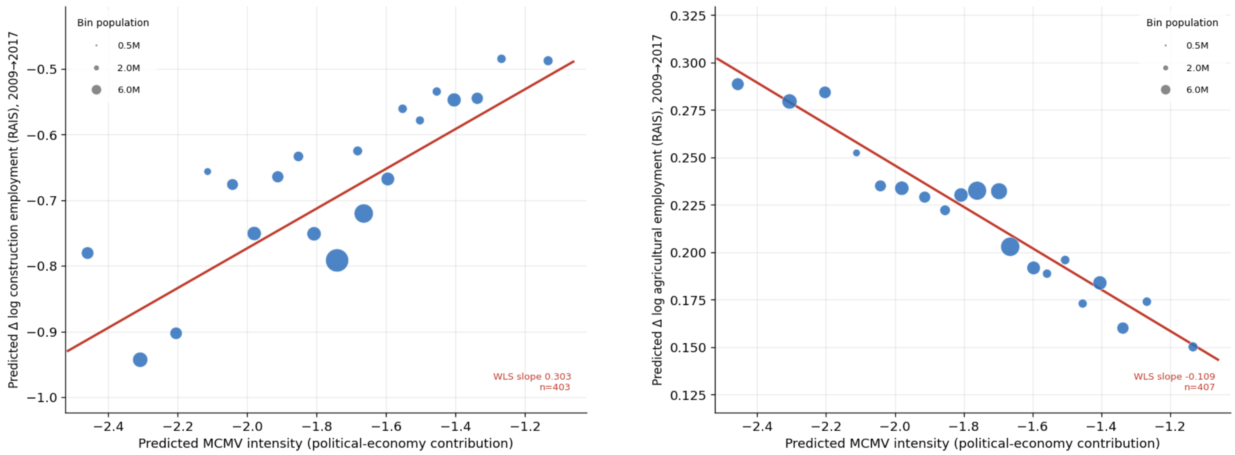


(a) Construction / real estate

(b) Manufacturing

Notes: Epicenter vs. never-entered, cumulative `did_multiplegt_dyn` event study (`placebo(5) effects(10)`, `trends_lin`, `only_never_switchers`); SEs clustered at the MMC level; event years -5 to $+10$; the horizontal line marks zero, with the treated (epicenter) side restricted to a single entry sector against the same never-entered controls.

Figure A7: Predicted RAIS formal-employment changes vs. predicted MCMV intensity.

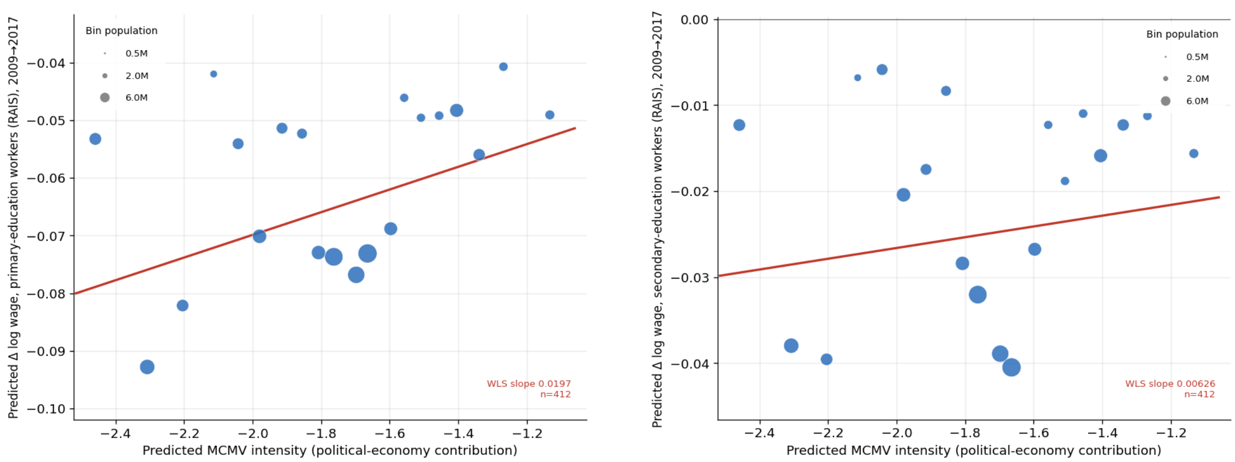


(a) Δ log construction employment

(b) Δ log agricultural employment

Notes: This figure shows binned scatters of MCMV's *predicted* housing program intensity, at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and *predicted* 2009-2017 changes in RAIS formal employment outcomes, using the same predictors. Both axes use the same set of political economy predictors and control for the same set of baseline socio-economic controls, defined in Section 2.4.2. Binned means are population-weighted; red line is weighted-least squares fit over $x \leq$99th percentiles; markers are proportional to bin population.

Figure A8: Predicted RAIS formal-wage changes vs. predicted MCMV intensity

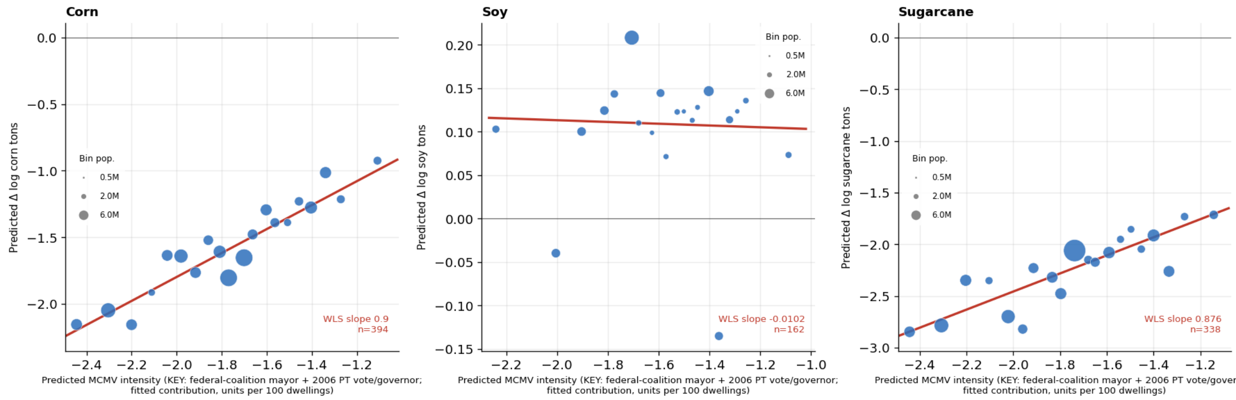


(a) Δ log wage, primary education

(b) Δ log wage, secondary education

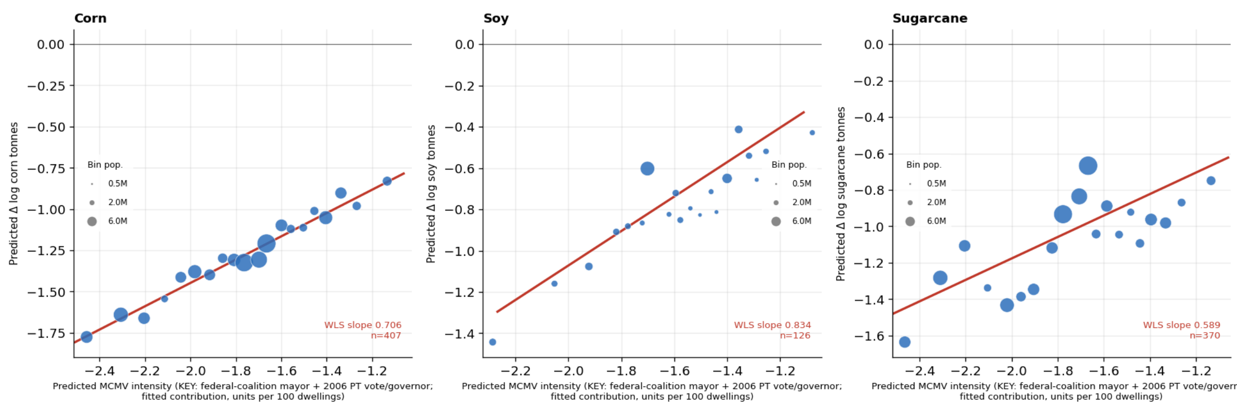
Notes: This figure shows binned scatters of MCMV's *predicted* housing program intensity, at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and *predicted* 2009-2017 changes in RAIS formal employment outcomes, using the same predictors. Both axes use the same set of political economy predictors and control for the same set of baseline socio-economic controls, defined in Section 2.4.2. Binned means are population-weighted; red line is weighted-least squares fit over $x \leq$99th percentiles; markers are proportional to bin population.

Figure A9: Predicted PAM crop output vs. predicted MCMV intensity



Notes: This figure shows binned scatters of MCMV's *predicted* housing program intensity, at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and *predicted* 2006-2017 changes in agricultural output per agricultural worker, where the former is measured from the PAM dataset separately by crop for Brazil's main temporary crops (soy, sugar, corn) and the latter is from the Agricultural census. Both axes use the same set of political economy predictors and control for the same set of baseline socio-economic controls, defined in Section 2.4.2. Binned means are population-weighted; red line is weighted-least squares fit over $x \leq$99th percentiles; markers are proportional to bin population.

Figure A10: Predicted Agricultural-Census crop output vs. predicted MCMV intensity



Notes: This figure shows binned scatters of MCMV's *predicted* housing program intensity, at the microregion level, measured as contracted units per 100 dwellings between 2009 and 2014, and *predicted* 2006-2017 changes in agricultural output per agricultural worker, where the former is measured from the Agricultural Census separately by crop for Brazil's main temporary crops (soy, sugar, corn) and the latter is from the Agricultural census. Both axes use the same set of political economy predictors and control for the same set of baseline socio-economic controls, defined in Section 2.4.2. Binned means are population-weighted; red line is weighted-least squares fit over $x \leq$99th percentiles; markers are proportional to bin population.

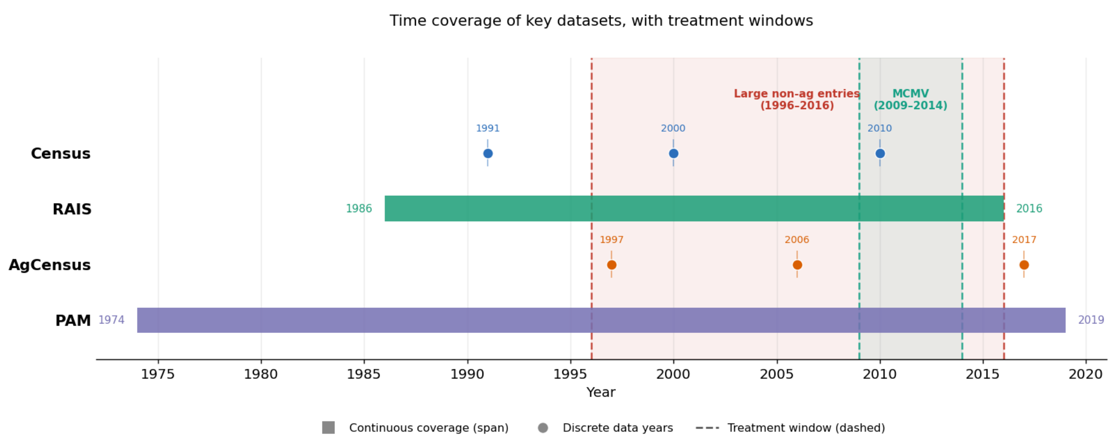
B Data Appendix

This appendix documents the data sources, sample construction, variable definitions, and deflation procedures underlying the analysis, collecting in one place the construction and cleaning decisions referenced in the main text.

B.1 Data sources

The datasets differ in frequency and coverage: RAIS is annual (1986–2016), agricultural outcomes are observed only at the Agricultural Census rounds (1997, 2006, 2017), and the Demographic Census (1991, 2000, 2010) supplies informality and migration. Figure B1 places every dataset and treatment window on a common timeline.

Figure B1: Data-coverage timeline with treatment windows



Notes: Continuous datasets are horizontal bars (RAIS 1986–2016; PAM 1974–2019); discrete-wave datasets are year markers (Demographic Census 1991/2000/2010; Agricultural Census 1997/2006/2017). Treatment windows: large firm entries 1998–2017; MCMV 2009–2014; SICONV 2008–2017.

RAIS (Relação Anual de Informações Sociais). Confidential linked employer–employee administrative records covering nearly the universe of formal employment, compiled annually by the Brazilian Ministry of Labour; we use the cleaned extracts of Engbom and Moser (2022) (1986–2016). Each employment relationship (*vínculo*) reports establishment, municipality, sector, contracted hours, schooling, demographics, and earnings (in multiples of the minimum wage). We aggregate to the municipality (MCA) level.

Agricultural Census (Censo Agropecuário, IBGE). Municipality-level agricultural data for 1995–96, 2006, and 2017 (we label 1995–96 as 1997 to align with the RAIS baseline; see Section B.5): farm area, agricultural labor, value of output, tractors, input expenditure shares, establishment counts by size class, and land tenure (rented vs. owned).

Demographic Census (Censo Demográfico, IBGE). Decennial individual microdata (1991, 2000, 2010) aggregated to the MCA: formal and informal employment by sector and schooling, and internal migration. The worker sample is restricted to individuals older than 10 who worked during the reference year.

PAM (Pesquisa Agrícola Municipal, IBGE). Annual municipal crop-production survey (1974–2019): quantity, harvested area, and value for the main temporary and permanent crops. Used for agricultural output per worker and crop-level responses (corn, soy, sugarcane, coffee).

MCMV (Minha Casa Minha Vida). Spatial allocation of MCMV projects, measured as units contracted 2009–2014; the raw municipal file records units by income tier (*Faixa*) and contracting window. Partisan and program-design allocation variables follow Simoni Junior and Dias (2025). See the Institutional Details Appendix for program detail.

SICONV (Sistema de Gestão de Convênios e Contratos de Repasse / Plataforma +Brasil). Records of voluntary (discretionary) federal-to-municipal transfers (*convênio / contrato de repasse*), 2008–2017. We measure intensity as committed *repasse* per capita and classify each agreement by project-text keywords (urban housing/sanitation, roads/paving, rural infrastructure). Downloaded from the federal open-data repository (https://repositorio.dados.gov.br/seges/detru/; Ministério do Planejamento / SEGES); *proposta* (municipality IBGE code) is joined to *convênio* (transferred amount) and aggregated to the MCA.

Terrain steepness. Slope rasters derived from the WorldClim v2.1 digital elevation model at 1 km^2 resolution (degrees, 0° flat to 90° vertical), made available for Brazil by INPE, intersected with the 2023 IBGE municipal-boundary shapefile. For each municipality we compute an area-weighted mean slope (cells weighted by the share of their area inside the municipality) and `prop_flat_under6`, the share of municipal area below 6°—the machinery-suitability cutoff of Nunn & Puga (2012, *REStat* 94(1)), after FAO (1993). Municipal measures are aggregated to the microregion (MMC) as population-weighted means.

Deflators and auxiliary data. Nominal values are deflated to 2017 reais with the IPCA (IBGE's official consumer price index; IBGE/SIDRA), using the real minimum-wage series from IPEADATA for RAIS earnings (Section B.4). Metropolitan-region indicators are from IBGE/`geobr`; MCA and MMC crosswalks are built from IBGE territorial-evolution records; and 1990–1994 trade-liberalization exposure, a balancing covariate, is from Dix-Carneiro and Kovak (2017).

B.2 Geographic units

The analysis runs at the Minimum Comparable Area (MCA) level and, for the agricultural and program-intensity designs, at the microregion (MMC) level; every MCA nests in one MMC, and we refer to MCAs as "municipalities." MCAs are the smallest time-consistent aggregations of municipalities: there are 5,352 MCAs covering 5,507 municipalities over 1997–2017, built from IBGE records on municipal splits and mergers between 1991 and 2010.[9]

B.3 Measurement of key variables

RAIS earnings and wages. From RAIS we aggregate formal employment and average schooling by sector, and two distinct compensation measures, both recorded in minimum-wage multiples and converted to 2017 reais (Section B.4): (i) *average monthly earnings* (`earn_mean_realw`), the headline compensation outcome in the event studies ("log average real monthly earnings," in the pooled, by-sector, and by-cohort results); and (ii) *wage per hour* (`per_hour_realw/realw_avwage`), monthly earnings $\times 12$ divided by annual hours worked, used in the by-education results. Both are employment-weighted means, rebuilt at the MCA as summed wage bill over summed employment.[10]

Agricultural variables. The agricultural wage is nominal agricultural labor expenses divided by agricultural workers (deflated to 2017 reais). Agricultural labor is from the Agricultural Census (RAIS measures only its formal component); RAIS estimates are matched to the census timing using annual or decade-preceding data for 1997, 2006, and 2016. Farms are classified by area as small ($<$ 100 ha), medium (100–500 ha), or large ($>$ 500 ha); average farm size is hectares per establishment; tractors per worker is tractors divided by agricultural workers; land tenure splits used land into rented and owned. Output per worker is total PAM crop value (`perm_valor_producao` + `temp_valor_producao`, deflated to 2017 reais) summed to the microregion, divided by Agricultural-Census workers; crop-level physical output (tons) is from PAM (and, in a robustness version, the Agricultural Census).

Informality and migration. Following Ulyssea (2018), informality is the share of employed workers aged 18–64 with non-zero earnings lacking a signed *Carteira de Trabalho*, from the Demographic Census, with each outcome decade matched to its baseline census (2010 for 2011–2016, 2000 for 2001–2010, 1991 for earlier). The non-agricultural informal share used in the heterogeneity analysis is the weighted ratio of informal to total non-agricultural employment among employed adults 18–64, with agriculture defined as CNAE divisions 1, 2, and 5. Internal-migration measures are built from the Demographic Census by socio-economic characteristics (1991 matched to 1995–97, 2000 to 2006, 2010 to 2017).

MCMV and SICONV intensity. Observed MCMV intensity is units contracted 2009–2014 per 100 baseline dwellings. Because units are not randomly allocated, predicted intensity is the fitted contribution

[9] IBGE, *Evolução da divisão territorial do Brasil* (https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/15771-evolucao-da-divisao-territorial-do-brasil.html). Of the 5,507 municipalities observed over 1997–2017, 1,016 split off from or joined others during the period.

[10] Resolving the question flagged in the main text: the RAIS *earnings* outcome in the event-study figures and ATT tables is average real *monthly earnings* (`earn_mean_realw`); the per-hour wage (`realw_avwage`) enters only the by-education panels. Confirmed against the RAIS difference-in-differences write-up.

of four political-economy predictors—a federal-coalition mayor in 2009, the core and swing PT presidential vote shares (1994–2006), and a PT state-governor indicator (2006)—from a regression that also includes 2000-vintage socioeconomic controls (urban-population share, a below-50,000-inhabitants indicator, a 2005 metropolitan-region indicator, and the agricultural, manufacturing, and formal employment shares) and macro-region fixed effects, keeping only the political-economy component; predicted outcome changes are built analogously. SICONV intensity is committed *repasse* per capita (2008–2017), with predicted intensity using the same predictors. (Allocation rules are detailed in the Institutional Details Appendix.)

Terrain measures. The heterogeneity analysis uses the standardized population-weighted mean municipal slope and `prop_flat_under6` (share of municipal area below 6°), both aggregated to the microregion (construction in Section B.1).

B.4 Deflation of nominal values

All nominal values are expressed in constant 2017 reais using the IPCA. RAIS earnings, recorded in minimum-wage multiples, are converted by multiplying by the real minimum-wage series (constant August-2017 reais) from IPEADATA. Agricultural-Census values (e.g. labor expenses) and PAM crop values are deflated with IPCA factors of 0.292 (1997), 0.532 (2006), and 1 (2017, the base year). Wherever a figure or table reports "real" values or "2017 reais," this deflation applies.

B.5 Firm-entry sample and treatment groups

A qualifying large entry is a new non-agricultural, non-public establishment with $\geq$200 employees at entry and size exceeding 0.1% of its microregion's baseline population, outside the agricultural-demand sectors (food and beverages, textiles, rubber and tobacco). Entries are restricted to 1998–2017 (so the 1997 Agricultural Census is a clean baseline) and grouped into Period 1 (1998–2006) and Period 2 (2007–2017), and classified as construction/real estate, manufacturing (IBGE transformation industries), or other (transport and communication, hospitality, mining, wholesale trade).

Analyses run at the MCA level while treatment is identified at the MMC level. An *epicenter* MCA contains a municipality with a qualifying entry (entry year = earliest qualifying entry; MMC entry year = earliest in the MMC); a *surrounding* MCA lies in the same MMC as an epicenter; and a *never-treated* MCA lies in an MMC with no large entry of any kind. A microregion is dropped from the never-treated pool if it had any large entry ($\geq$100 employees or $>$0.1% of baseline population) in any sector, and dropped from the sample entirely if it had a large entry of any sector before 1998. The sample comprises 81 epicenter municipalities (69 microregions; 94 entries), 23 entering in Period 1 and 58 in Period 2.

C Institutional Details: MCMV and SICONV

This appendix gives institutional background on the two federal construction-related programs used in the analysis. Section C.1 covers MCMV's legal basis, allocation rules, income tiers, and contracting; and Section C.2 covers the SICONV transfer system.

C.1 Minha Casa Minha Vida (MCMV)

Overview and intensity measure. MCMV is Brazil's flagship federal low-income housing program, launched in March 2009. It subsidizes construction across income tiers (*Faixas*), with the lowest, *Faixa 1*, most heavily subsidized and targeted at the poorest households. We measure intensity as units contracted 2009–2014 per 100 baseline dwellings (predicted-intensity construction is in the Data Appendix, Section B.3).

Legal basis and allocation rule. Established by Lei n. 11.977/2009 (regulating Decreto n. 6.962/2009, Art. 5), the program distributes federal resources across states *by estimated state housing deficit* (IBGE/PNAD 2007), while the Ministry of Cities retains discretion to reallocate *within* states by "capacidade de contratação" (the existence of qualified projects). Allocation is further conditioned on *population size* (the FAR modality in larger cities, dedicated offerings for municipalities under 50,000 inhabitants, per-municipality ceilings) and *urban location* (serviced urban land; priority to metropolitan regions and state capitals in the higher tiers). These rules motivate the housing-deficit, below-50,000-inhabitants, urban-share, and metropolitan-region predictors.

Faixa 1 / FAR modality. Faixa 1 units are produced through the Fundo de Arrendamento Residencial (FAR): the federal government funds construction, while municipalities and states facilitate projects by donating regularized, well-located public land and providing tax relief or other support. Because the relevant concept is *buildable urban* land—not captured by the available predetermined measure of government-owned *agricultural* land—public-land availability is not used as a predictor.

Partisan allocation and contracting. Within the law's within-state discretion, Simoni Junior and Dias (2025) show that 2009 mayoral alignment with the federal governing coalition strongly predicts Faixa 1 contracts over 2009–2014, prior to the 2010 presidential election; this is the basis for the political-economy predictors (federal-coalition mayor 2009; core and swing PT presidential vote 1994–2006; PT state-governor 2006). FAR units were contracted on a continuous, rolling basis (the "balcão" model)—proposals analyzed and contracted by Caixa in order of completeness, against per-state targets—so contracting accumulates over the window rather than at a single date. On the electoral calendar: mayoral elections fell in October 2008 and 2012 (terms from January), the presidential/gubernatorial election in October 2010, and the next presidential election in October 2014; the 2009–2012 window coincides with the Lula presidency and first mayoral mandate, and the 2011–2014 window spans the start of the Dilma presidency and second mandate.[11]

C.2 SICONV (discretionary federal transfers)

Overview. SICONV (Sistema de Gestão de Convênios e Contratos de Repasse) records voluntary (discretionary) federal-to-municipal transfers through agreements (*convênios / contratos de repasse*) funding urban infrastructure, paving, sanitation, housing, and rural development. Because these transfers are discretionary rather than formula-based, political-economy predictors are informative about their allocation, so SICONV serves as a second source of construction-investment variation analyzed with the same predicted-on-predicted design as MCMV.

Role and measurement. SICONV transfers show the same agricultural gradients as MCMV, concentrated in the urban-housing/sanitation category—not rural infrastructure, which shows no relationship. Because predicted MCMV and SICONV intensities are built from the same predictors and are correlated, the SICONV results mainly confirm that the agricultural reorganization reflects construction-driven urban labor demand broadly, not MCMV's specific rules. We measure intensity as committed *repasse* per capita over 2008–2017 and classify each agreement by project-text keywords (urban housing/sanitation, roads/paving, rural infrastructure). Data are from the federal open-data repository (Ministério do Planejamento / SEGES; https://repositorio.dados.gov.br/seges/detru/); join and aggregation details are in the Data Appendix (Section B.1).

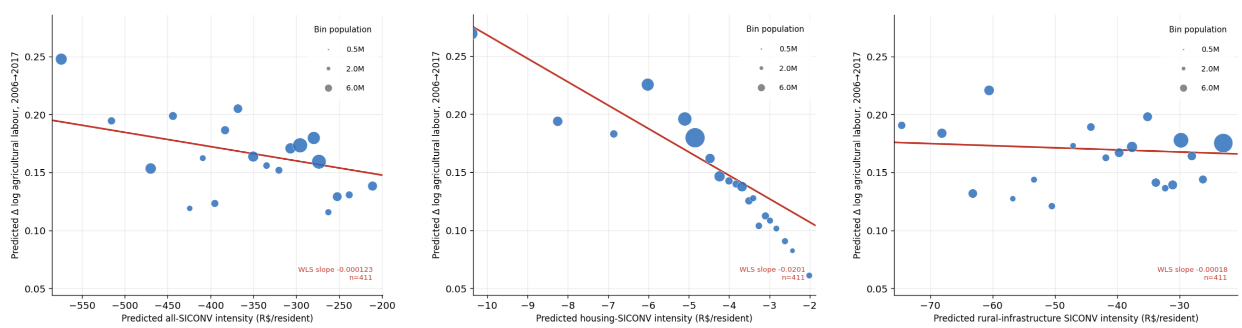
[11]Sources: Lei n. 11.977/2009 (https://www2.camara.leg.br/legin/fed/lei/2009/lei-11977-7-julho-2009-589206-publicacaooriginal-114190-pl.html); Decreto n. 6.962/2009; Ministério das Cidades and Caixa documentation on the FAR/Faixa 1 modality and "balcão" contracting; Simoni Junior and Dias (2025) for the legal criterion and partisan-electoral data.

Table C1: SICONV investment categories, classified from project text descriptions (2008–2017)

Investment category (text-keyword)	Total repasse (R$ bn)	Share (%)	Convênios (n)
Construction / works (any)	54.71	58.6	67,242
Urban housing + sanitation	14.94	16.0	12,556
Roads / paving	14.73	15.8	26,658
Rural infrastructure	7.11	7.6	14,444
Keyword components (subsets of the groups above)			
Housing (*habitação*)	1.02	1.1	1,160
Sanitation (*saneamento*)	11.20	12.0	7,904
Urbanization (*urbanização*)	3.27	3.5	3,820
Paving (*pavimentação*)	14.73	15.8	26,658
Rural (*rural*)	7.11	7.6	14,444
All SICONV (2008–2017)	93.30	100.0	127,780

Notes: Committed SICONV *repasse* 2008–2017, classified into investment categories by keywords in each agreement's (*convênio / contrato de repasse*) project description. "Construction / works (any)" is the union of the urban-housing/sanitation, roads/paving, and rural-infrastructure groups; the keyword components listed below are overlapping subsets. Total *repasse* in billions of R$, with the share of all SICONV and the number of agreements (n).

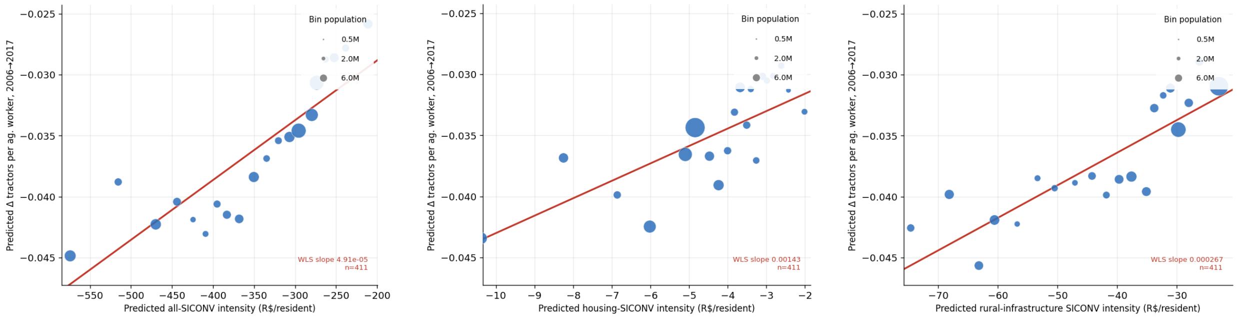
Figure C1: Predicted Δ log agricultural labor vs. predicted SICONV intensity



(a) All SICONV transfers (b) Housing transfers only (c) Rural infrastructure

Notes: Microregion level; both axes are the political-economy predicted contribution. Population-weighted binned scatter (20 bins) with WLS fits, common y-axis. Change measured 2006→2017.

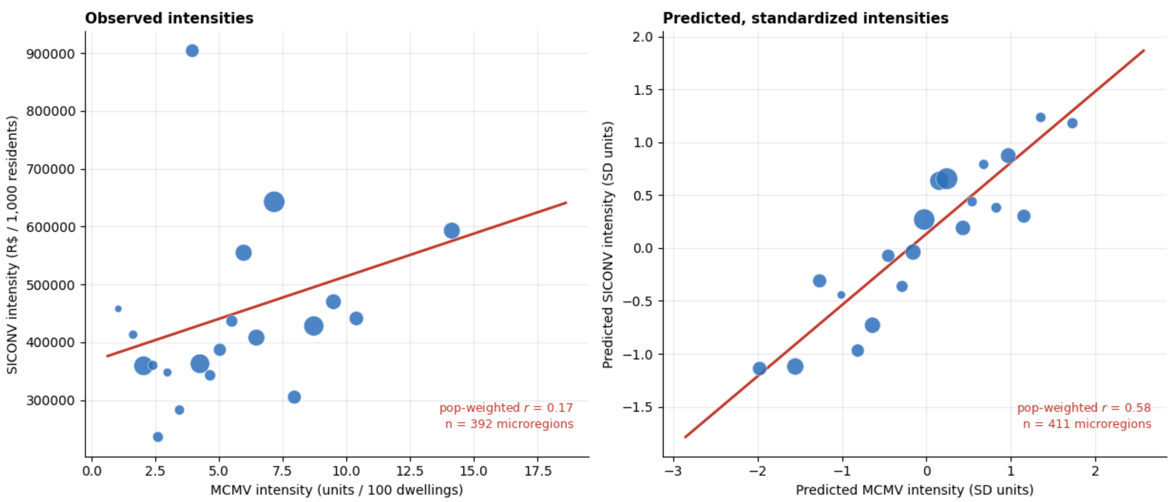
Figure C2: Predicted Δ tractors per worker vs. predicted SICONV intensity



(a) All SICONV transfers (b) Housing transfers only (c) Rural infrastructure

Notes: Microregion level; both axes predicted; population-weighted binned scatter (20 bins) with WLS fits, common y-axis. Change measured 2006→2017.

Figure C3: MCMV vs. SICONV intensity across microregions



Notes: Microregion level; population-weighted binned scatter (20 equal-count bins). Left: observed intensities; right: predicted, standardized intensities. Red line: WLS fit; population-weighted.

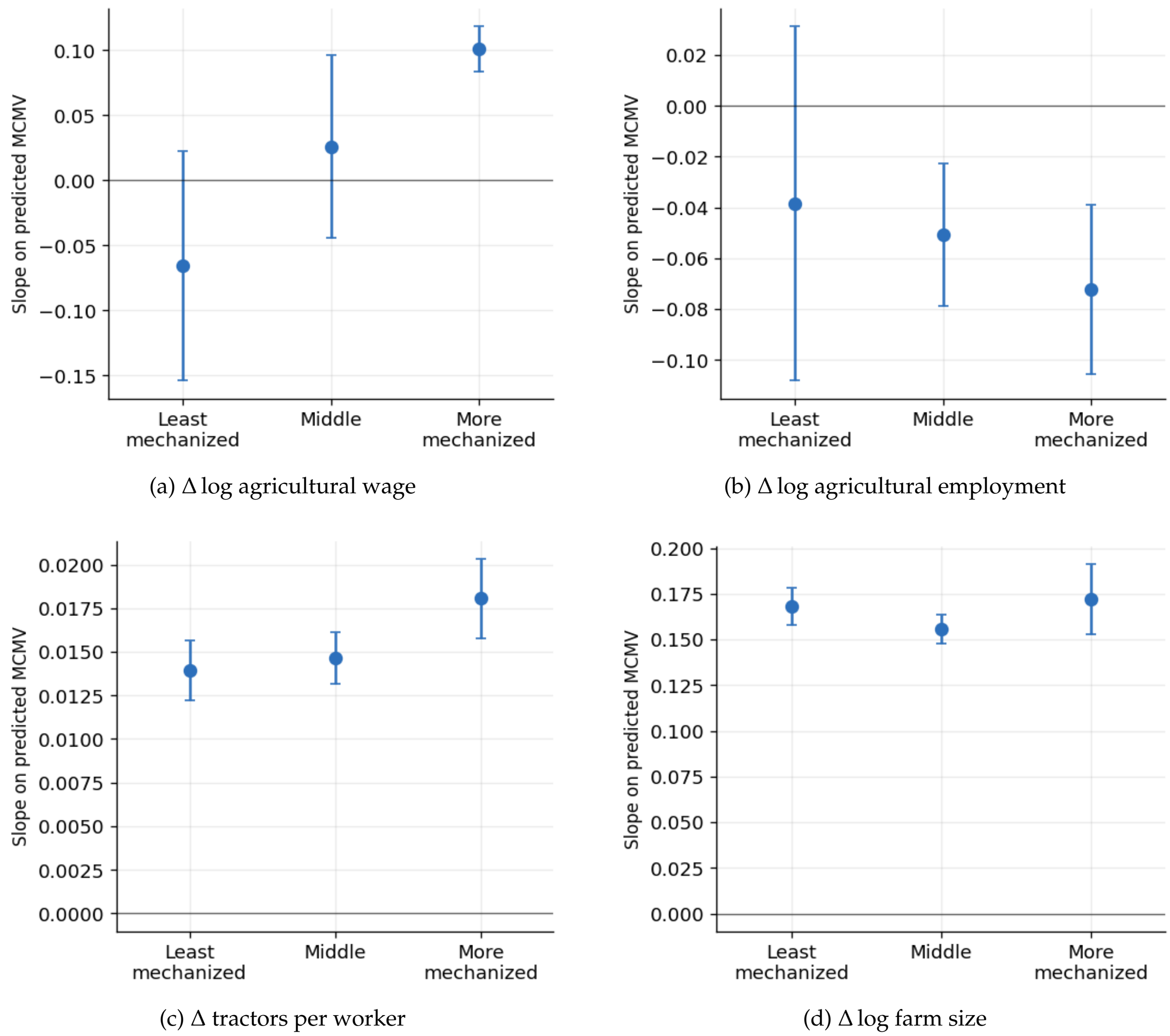
D Additional results on heterogeneity of the MCMV gradient

Table D1: Predicted outcome changes on predicted MCMV intensity × heterogeneity dimension

	Terrain		Baseline characteristic	
	(1) Mean slope (z)	(2) Prop. flat ($< 6^\circ$)	(3) Non-ag inf. 2000	(4) Tractors/wkr 2006
Panel A: Δ log agricultural wage				
Predicted MCMV	0.0825***	0.1346***	0.0443*	0.0220
	(0.0273)	(0.0321)	(0.0229)	(0.0253)
× dimension	0.0132	-0.0684	-0.1086***	0.0406*
	(0.0152)	(0.0561)	(0.0273)	(0.0231)
Dimension main effect	0.0487	-0.2149*	-0.2308***	0.1139***
	(0.0297)	(0.1128)	(0.0483)	(0.0385)
Observations	411	411	411	411
Panel B: Δ log agricultural employment				
Predicted MCMV	-0.0550***	-0.1126***	-0.0419***	-0.0531***
	(0.0173)	(0.0266)	(0.0129)	(0.0171)
× dimension	-0.0194*	0.0818**	0.0584***	-0.0288
	(0.0113)	(0.0413)	(0.0139)	(0.0181)
Dimension main effect	-0.0352*	0.1485**	0.0897***	-0.0505*
	(0.0192)	(0.0702)	(0.0245)	(0.0287)
Observations	411	411	411	411
Panel C: Δ tractors per worker				
Predicted MCMV	0.0172***	0.0179***	0.0162***	0.0151***
	(0.0007)	(0.0015)	(0.0006)	(0.0006)
× dimension	-0.0000	-0.0008	-0.0023***	0.0005
	(0.0005)	(0.0019)	(0.0008)	(0.0007)
Dimension main effect	0.0004	-0.0035	-0.0056***	0.0027**
	(0.0009)	(0.0036)	(0.0015)	(0.0011)
Observations	411	411	411	411
Panel D: Δ log farm size				
Predicted MCMV	0.1663***	0.1679***	0.1642***	0.1612***
	(0.0041)	(0.0111)	(0.0032)	(0.0042)
× dimension	-0.0008	-0.0021	-0.0038	-0.0008
	(0.0035)	(0.0142)	(0.0045)	(0.0054)
Dimension main effect	-0.0026	-0.0019	-0.0097	0.0053
	(0.0060)	(0.0247)	(0.0099)	(0.0078)
Observations	411	411	411	411

Notes: Microregion level. Columns (1)–(2) interact predicted MCMV with the two terrain measures (standardized mean municipal slope; the share of municipal area below 6°); columns (3)–(4) with the standardized baseline characteristic (2000 non-agricultural informal-employment share, 2006 tractors per worker). Column (4) additionally controls for individual 4th-degree polynomials in terrain slope and baseline farm size (not shown). No second-stage region FE; 2000-population weights; microregion-clustered SE; $N = 411$. $^{*}/^{**}/^{***}$: $p < 0.10/0.05/0.01$.

Figure D1: Predicted outcome changes vs. predicted MCMV intensity, by baseline mechanization tercile.



(a) Δ log agricultural wage

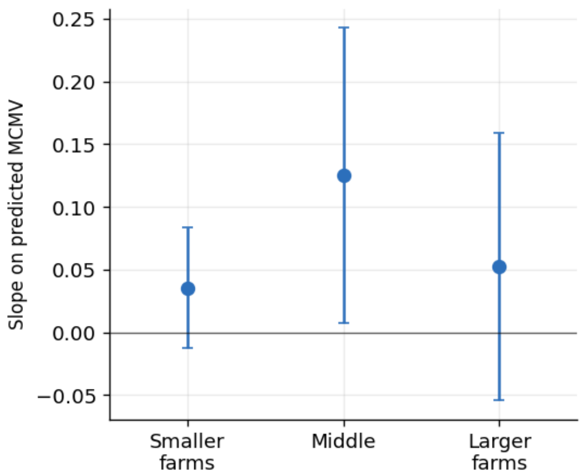
(b) Δ log agricultural employment

(c) Δ tractors per worker

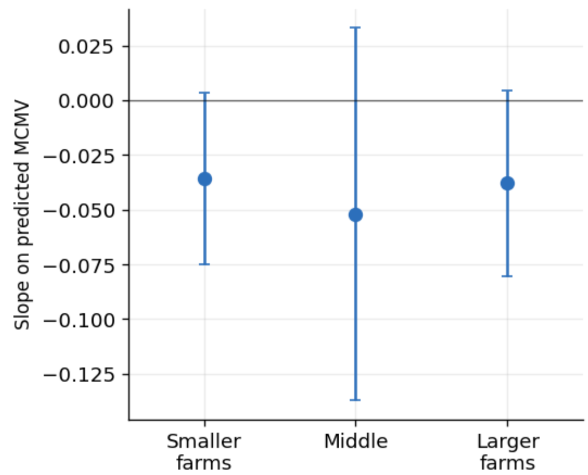
(d) Δ log farm size

Notes: Microregion level; predicted-on-predicted fits within terciles of 2006 tractors per agricultural worker (Least / Middle / More mechanized; microregions with no tractors retained at zero). No second-stage region fixed effects; SE clustered at the microregion. Tercile cutoffs are printed at the bottom of each panel; continuous interactions in Table D1.

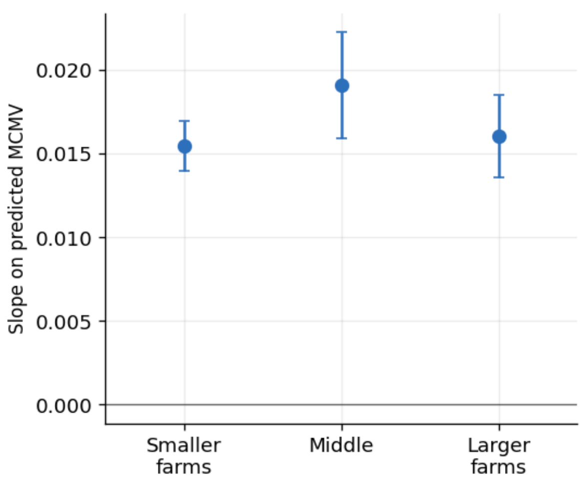
Figure D2: Predicted outcome changes vs. predicted MCMV intensity, by baseline farm-size tercile.



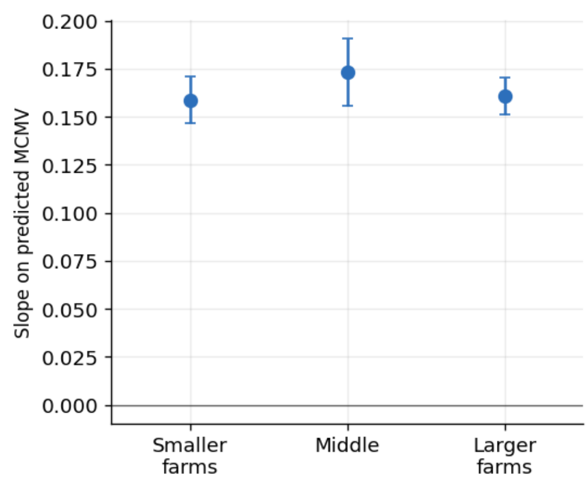
(a) Δ log agricultural wage



(b) Δ log agricultural employment



(c) Δ tractors per worker



(d) Δ log farm size

Notes: Microregion level; predicted-on-predicted fits within terciles of 2006 average farm size, hectares per establishment (Smaller / Middle / Larger farms). No second-stage region fixed effects; SE clustered at the microregion. Tercile cutoffs are printed at the bottom of each panel; continuous interactions in Table D1.

E Model Derivations and Proofs

E.1 Derivation of Reduced-Form Production Function from Tasks

We derive equation (11) from the task-based production technology in (9)–(10).

Optimality conditions and equal expenditure shares. The planner's first-order condition for task output $x_j(b)$ is

$$\lambda_a \gamma \frac{Y_{aj}}{x_j(b)} = p_x(b), \tag{36}$$

where $p_x(b)$ is the shadow cost of a unit of task output. From equation (36), the expenditure on each task is equalized:

$$p_x(b)\, x_j(b) = \lambda_a \, \gamma \, Y_{aj} \quad \text{for all } b \in [0,1]. \tag{37}$$

Total task expenditure is $\int_0^1 p_x(b)\, x_j(b)\, db = \lambda_a \gamma Y_{aj}$, consistent with γ being the output elasticity of the composite input.

Input choice within each task. For a non-mechanizable task $b < \tilde{\beta}$ (only labor feasible), the optimality condition is

$$p_x(b)\, a_n(b) = \mu_N, \tag{38}$$

so $n_j(b) = \lambda_a \gamma Y_{aj}/\mu_N$ for all $b < \beta$ (equal labor per labor task). For a mechanizable task $b \geq \tilde{\beta}$ performed with capital (i.e. $b \geq \beta$), the optimality condition is

$$p_x(b)\, a_k(b) = \lambda_k, \tag{39}$$

so $k_j(b) = \lambda_a \gamma Y_{aj}/\lambda_k$ for all capital tasks (equal capital per capital task). The threshold activity $\beta \in [\tilde{\beta}, 1]$ is indifferent:

$$\frac{a_n(\beta)}{a_k(\beta)} = \frac{\mu_N}{\lambda_k}. \tag{40}$$

Aggregation. Let $N_{aj} = \int_0^\beta n_j(b)\, db$ and $K_{aj} = \int_\beta^1 k_j(b)\, db$. From the equal-input conditions above, $n_j(b) = N_{aj}/\beta$ for all $b < \beta$ and $k_j(b) = K_{aj}/(1-\beta)$ for all $b \geq \beta$. Substituting into the production function:

$$\begin{aligned} Y_{aj} &= Z_{aj} \exp\!\left(\gamma \left[\int_0^\beta \ln\!\left(a_n(b) \frac{N_{aj}}{\beta} \right) db + \int_\beta^1 \ln\!\left(a_k(b) \frac{K_{aj}}{1-\beta} \right) db \right] \right) L_{aj}^{\gamma_L} \\ &= Z_{aj}\, A(\beta)\, \left(K_{aj}^{1-\beta} N_{aj}^{\beta}\right)^{\gamma} L_{aj}^{\gamma_L}, \end{aligned} \tag{41}$$

where

$$A(\beta) = \exp\!\left(\gamma\left[\int_0^{\beta} \ln a_n(b)\, db + \int_{\beta}^{1} \ln a_k(b)\, db\right]\right) \beta^{-\gamma\beta}(1-\beta)^{-\gamma(1-\beta)}. \tag{42}$$

This is the reduced-form production function in equation (11). Note that $A(\beta)$ depends only on β (the actual task split), not on $\tilde{\beta}$ (the mechanizable threshold): the efficiency of the production arrangement depends on which tasks use which input, not on how many could in principle use capital.

E.2 Proof of Proposition 2

Setup. Suppose the economy is in a stationary equilibrium with factor prices (μ_N^0, λ_k^0) and the farm's mechanizable threshold is $\tilde{\beta}_0$, satisfying (14). The actual mechanization threshold is $\beta_0 \geq \tilde{\beta}_0$, satisfying (40). We consider a farm that is at its mechanizable limit, $\beta_0 = \tilde{\beta}_0$. To avoid the notation confusion, we introduce separate symbols: $\tilde{\beta}_0$ is the pre-shock *mechanizable* threshold (and the pre-shock *actual* threshold since the farm is constrained), and $\beta_0^* \equiv \beta^*(\mu_N^0/\lambda_k^0)$ is the frictionless optimum. By assumption, $\beta_0^* < \tilde{\beta}_0 = \beta_0$ (the farm would like to mechanize further but its mechanizable set does not reach no cost optimum).

Step 1: Temporary shock. In period T, wages rise temporarily to $\mu_N^1 > \mu_N^0$, with λ_k unchanged. The frictionless optimum β^* falls (equation (12) has a unique solution decreasing in μ_N/λ_k). Because $\beta^*(\mu_N^1) < \tilde{\beta}_0$, the farm cannot achieve the frictionless optimum without expanding the mechanizable set. Since the net gain from paying I and lowering $\tilde{\beta}$ is positive for μ_N^1 large enough, the farm pays I and sets $\tilde{\beta}' = \beta^*(\mu_N^1)$.

Step 2: Return to original wages. In period $T+1$, wages return to μ_N^0. The no cost optimum reverts to $\beta_0^* = \beta^*(\mu_N^0/\lambda_k^0)$, which by the setup satisfies $\tilde{\beta}' < \beta_0^* < \tilde{\beta}_0 = \beta_0$.

Step 3: Post-shock equilibrium. Under the irreversibility (15), the mechanizable threshold remains at $'$: no cost is required to keep the expanded set. The farm can now freely set $\beta \in [\tilde{\beta}', 1]$. Since $\tilde{\beta}' < \beta_0^* < \tilde{\beta}_0$, the no cost optimum β_0^* is now freely attainable: the farm sets $\beta = \beta_0^*$ from condition (40) at the original prices. Crucially, this is *strictly more mechanization* than the pre-shock equilibrium: before the shock, the farm was constrained to $\beta = \tilde{\beta}_0 > \beta_0^*$ and could not reach its frictionless optimum; after the shock, the expanded mechanizable set allows it to do so at zero additional cost. The temporary wage increase has therefore generated a permanent improvement in actual mechanization (β falls permanently from $\tilde{\beta}_0$ to β_0^*), in addition to expanding the mechanizable set ($\tilde{\beta}$ falls permanently from $\tilde{\beta}_0$ to $\tilde{\beta}'$).

In terms of actual capital-to-labor ratios, if $\beta_0 > \tilde{\beta}'$, the farm operates with the same β as before but with a larger mechanizable set. Any further wage increase now requires a smaller additional fixed cost to reach a new $\tilde{\beta}'' < \tilde{\beta}'$, since the gap between β_0 and $\tilde{\beta}''$ is smaller. Combined with the farm-exit and land-consolidation channel (Proposition 3), this effect is further amplified in general equilibrium: surviving farms are larger, reducing the per-task fixed cost $I/\bar{k}$ and making additional mechanization more attractive. The two channels therefore reinforce each other. In the aggregate, farms that have paid the fixed cost during the boom are therefore more responsive to future wage increases—the model generates not only persistence but also increasing returns to temporary shocks in terms of subsequent mechanization. □

F Full Equilibrium Conditions

Since the above economy is efficient, we can compute allocations using the planner's problem (stationary equilibrium).

Entry conditions. Let $\psi_{n,t}$ denote the planner's shadow value of a unit mass of non-agricultural firms—the Lagrange multiplier on the firm-measure evolution constraint (22). The old form $\mu_{N,t}(1 - G(\bar{Z}_n)) = F_n \mu_{N,t}$ is dimensionally ill-posed: it conflates the shadow wage with the shadow value of a firm and collapses to the uninformative condition $1 - G(\bar{Z}_n) = F_n$, which is independent of factor prices and profits. The correct planner's optimality condition for the mass of entrants $m_{n,t}$ is:

$$\psi_{n,t+1}\big(1 - G(\bar{Z}_{n,t+1})\big) = F_n\, \mu_{N,t}. \tag{43}$$

The shadow value of an active firm satisfies the Bellman equation

$$\psi_{n,t} = \nu\big[\bar{\pi}_{n,t+1} + \psi_{n,t+1}(1 - \delta_{Mn})\big], \tag{44}$$

where $\bar{\pi}_{n,t+1} = \int_{\bar{Z}_n} (\lambda_n Y_{ni} - \mu_N N_{ni} - f_n \mu_N)\, g_n(Z)\, dZ$ is average net profit (valued at planner prices) in non-agriculture. Combining (43) and (44) in steady state yields the free-entry condition (50) below, which is unchanged.

Analogous conditions hold for agriculture, with $\psi_{a,t}$ replacing $\psi_{n,t}$ and capital rental and land costs entering $\bar{\pi}_a$.

Zero-profit conditions. At the entry thresholds, profits of the marginal firm are zero:

$$(1 - \alpha)\, \lambda_n\, \bar{Y}_n(Z_n) = f_n\, \mu_N, \tag{45}$$

$$(1 - \gamma - \gamma_L)\, \lambda_a\, \bar{Y}_a(Z_a) = f_a\, \mu_N, \tag{46}$$

which can be rewritten in terms of average-firm objects using the Pareto structure of G_n and G_a.

Capital allocation. The investment Euler equation gives

$$\lambda_{n,t} = \nu\left[\lambda_{n,t+1}(1 - \delta_k) + \lambda_{a,t+1}(1 - \beta)\gamma \frac{\bar{Y}_{a,t+1}}{K_{a,t+1}}\right]. \tag{47}$$

In the steady state, λ_n is constant and $(1 - \beta)\gamma\lambda_a \bar{Y}_a/(\lambda_n K_a) = \hat{\delta}_k/\nu$, where $\hat{\delta}_k = 1 - \nu(1 - \delta_k)$.

Relative labor allocations. Cross-farm (within sector) productivity differences determine relative employment through (35). The relative employment between the informal sector and non-agriculture is

$$\frac{Z_s}{Z_n}\left(\frac{N_s}{N_n}\right)^{\alpha-1} = \frac{\lambda_s}{\lambda_n}. \tag{48}$$

The relative labor productivity between agriculture and the informal sector is given by equation (34). These conditions, combined with goods-market clearing (24)–(25) and the household's Stone-Geary optimality conditions

$$\lambda_s = \frac{\theta_s}{Z_s N_s^\alpha - \bar{c}_s}, \qquad \lambda_n = \frac{\theta_n}{M_n(1-s)Z_n N_n^\alpha - \bar{c}_n}, \tag{49}$$

complete the characterization of the stationary equilibrium.

Free entry, steady state. Combining the entry condition, the value equation, and the zero-profit condition yields the steady-state free-entry condition for non-agriculture:

$$\bar{\pi}_n = \frac{\hat{\delta}_{Mn}\, F_n\, \mu_N}{\nu(1-G(\bar{Z}_n))}, \tag{50}$$

where $\hat{\delta}_{Mn} = 1-\nu(1-\delta_{Mn})$. An analogous condition holds for agriculture, with the capital and land costs entering $\bar{\pi}_a$ and with the farm's mechanization decision feeding back into average profits through β and $\tilde{\beta}$.

G Computational Algorithm

We solve the model by iterating on the shadow price of non-agricultural output λ_n (equivalently, the shadow wage μ_N) until the labor market clears. The algorithm proceeds in the following steps.

1. *Guess* λ_n and λ_a. Use the entry conditions and zero-profit conditions to solve for the productivity thresholds $\bar{Z}_n$ and $\bar{Z}_a$.

2. *Capital-output ratio.* Use the steady-state Euler equation (47) and the optimality condition for agricultural consumption to solve for the product $M_a K_a$:

$$1 = \nu \left[(1 - \delta_k) + \frac{\lambda_a (1 - \beta) \gamma \bar{Y}_a}{\lambda_n K_a} \right], \qquad \lambda_a = \frac{\theta_a}{M_a \bar{Y}_a - \bar{c}_a}.$$

3. *Investment.* Use the law of motion for capital in steady state to get $X = \delta_k M_a K_a$.

4. *Savings rate.* Use the household optimality condition for non-agricultural consumption to infer

$$s = X / (M_n \bar{Y}_n).$$

5. *Non-agricultural firm size.* From the free-entry condition and the zero-profit condition in non-agriculture,

$$N_n = \frac{\hat{\delta}_{Mn} F_n \, \alpha}{\nu (1 - G(\bar{Z}_n))(1 - \alpha)}.$$

6. *Shadow wage.* Infer μ_N from the optimality condition for labor in non-agriculture, $\mu_N = \alpha \lambda_n Z_n N_n^{\alpha - 1}$.

7. *Measure of agricultural farms.* Use the household optimality condition for agricultural consumption and the free-entry condition in agriculture to solve for M_a:

$$M_a = \frac{(\theta_a + \bar{c}_a \lambda_a)(1 - \gamma - \gamma_L)}{\hat{\delta}_{Ma} F_a \mu_N / (\nu (1 - G(\bar{Z}_a)))}.$$

8. *Capital stock.* Use $M_a K_a$ from step 2 and M_a from step 7 to recover the per-farm capital K_a.

9. *Measure of non-agricultural firms.* From the household optimality condition for non-agricultural consumption,

$$M_n = \frac{\theta_n + \bar{c}_n \lambda_n}{(1 - s) Z_n N_n^{\alpha}}.$$

10. *Agricultural employment.* Use the optimality conditions for labor in agriculture and in the informal sector to solve for N_a (the employment of the average active farm). With the production function (11),

the capital-to-labor ratio satisfies the first-order condition

$$(1-\beta)\gamma\frac{\lambda_a \bar{Y}_a}{K_a} = \lambda_k, \qquad \beta\gamma\frac{\lambda_a \bar{Y}_a}{N_a} = \mu_N.$$

11. *Informal sector prices.* From the household optimality condition in the informal sector,

$$\lambda_s = \frac{\theta_s}{Z_s N_s^\alpha - \bar{c}_s}.$$

12. *Update agricultural goods price.* From the household optimality condition,

$$\lambda_a = \frac{\theta_a}{M_a \bar{Y}_a - \bar{c}_a}.$$

13. *Labor market clearing.* Compute total labor demand:

$$\text{LHS} = M_n N_n + N_s + M_a N_a + \frac{\hat{\delta}_{Ma} M_a}{\nu(1-G(\bar{Z}_a))} F_a + \frac{\hat{\delta}_{Mn} M_n}{\nu(1-G(\bar{Z}_n))} F_n + f_a M_a + f_n M_n.$$

If LHS $> \bar{N}$, raise λ_n (and μ_N accordingly); if LHS $< \bar{N}$, lower λ_n. Iterate until convergence.

Mechanization in the algorithm. In step 10, β and $\tilde{\beta}$ are determined jointly with the employment and capital levels. For a given capital stock K_a and wage μ_N, the actual mechanization threshold β satisfies the task indifference condition $a_n(\beta)/a_k(\beta) = \mu_N/\lambda_k$ (equation (12) in the main text). The mechanizable threshold $\tilde{\beta}$ satisfies equation (14), using $\bar{k} = K_a/(1-\tilde{\beta})$; this requires a fixed-point iteration within step 10, using the current-iteration value of λ_a in equation (14). In the stationary equilibrium with the irreversibility constraint, $\tilde{\beta}$ is at its long-run lower bound, meaning the farm has exhausted all profitable mechanization investments at the prevailing factor prices.