The Origins of Structural Transformation

Julieta Caunedo * Mayara Felix † Kristina Manysheva ‡

September 9, 2025
Preliminary and in progress

Abstract

We study how labor market shocks originating in non-agriculture affect the organization of agricultural production. Using rich data from Brazil between 1986 and 2017, we show that the entry of large non-agricultural firms leads to persistent increases in local wages, declines in agricultural employment, and a shift toward more capital-intensive farming. Farms consolidate, the number of small operations declines, and mechanization increases. To study the magnitude of different channels contributing to this reorganization of production, we develop a general equilibrium model that features heterogeneous firms and farms through decreasing returns, an endogenous mechanization margin in agriculture and income effects. Farms optimally substitute capital for labor in response to rising wages following labor demand increase. At the same time, households demand relatively less agricultural products as their income raises with higher productivity in non-agriculture. Calibrated to match the Brazilian economy, our model predicts that a reduction in entry costs in non-agriculture–sufficient to generate the wage increase observed in the data–leads to labor reallocation out of agriculture, farm exit, and capital deepening. When mechanization is held fixed, these adjustments are substantially attenuated, highlighting the role of endogenous technology adoption. Finally, we use the model to explore the role of heterogeneity in initial mechanization to can rationalize disparate responses to similar non-agriculture labor demand shocks. We find labor reallocation is non-linear in baseline mechanization and responses are stronger in less mechanized markets.

Keywords: Structural Transformation, Agriculture, Firm Entry, Labor Markets.
JEL Codes: O13; O47, O57, Q12.

*Cornell University and CEPR, email: julieta.caunedo@cornell.edu,
†Yale University and NBER, email: mayara.felix@yale.edu,
‡Columbia University, email: km3924@columbia.edu.

We thank audiences at Columbia University, Princeton University, Barcelona Summer Forum, SED Winter '24, the Richmond Fed, and the Minneapolis Fed for their comments and suggestions. Guilherme Bayma, 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

The process of structural transformation—the reallocation of labor and resources from agriculture to manufacturing and subsequently to services—is a defining feature of economic development. Classical theories predict that as economies develop, differences in sectorial productivity trends and income effects lead the reallocation of labor into non-agricultural sectors (Herrendorf et al., 2014). Evidence from industrialized economies suggests that changes in industrial technology attracting labor out of agriculture were key to labor reallocation until the 1920s, while improvements in agricultural technology drove labor reallocation post 1960s (Alvarez-Cuadrado and Poschke, 2011). While either type of technological change resulted in productivity improvements in agriculture, the underlying forces driving those improvements differ, and so the channels driving labor reallocation and its implications for wages. The study of recent developers has put particular emphasis on technological change and shocks to agriculture (Bustos et al., 2016; Imbert and Ulyssea, 2023). In contrast, we know relatively little about the impact of positive shocks in non-agriculture for the organization of production in agriculture, despite their growing relevance in the context of industrial development, urbanization, and firm expansion.

In this paper, we study how entries of large-scale non-agricultural firms that shifts labor demand affect the organization of agricultural production, including changes in farm size, capital intensity, and labor engagement. Canonical models of structural change suggest that firm entry, and the associated improvements in non-agriculture productivity, leads to labor reallocation into agriculture (rather than away from it).[1] Combining rich panel data from Brazil's municipalities—-including agricultural censuses, population censuses, and employer-employee matched data for the overall economy–we show that non-agriculture firm entry is associated with labor reallocation out of agriculture and increases in wages. At the same time, non-agriculture entry is associated with more mechanization in agriculture and small farm exit. These effects are identified from staggered different and different design comparing municipalities that received large entry shocks, and those in the same micro-region that did not. We show that most of these entries are accounted for by the construction and the real-state sectors, and that those are, in turn, correlated with a government program for low-income housing.[2]

We develop a general equilibrium model in which the organization of agricultural production changes endogenously to relative prices, and therefore, developments in the non-agricultural sector. Decreasing returns in production technologies lead to a non-trivial size distribution of firms in non-agriculture and farms in agriculture (Hopenhayn, 1992) but nests, in its aggregate form, the properties of standard models of structural change. One key departure form a canonical model is that capital intensity in the agricultural sector is endogenous, Zeira (1998). As in Caunedo and Kala (2021), the model is tractable despite heterogeneity in farm productivity and potentially, incentives to mechanize. The measure of operating firms

[1] Herrendorf et al. (2014) provide a review of research on structural transformation.

[2] Our results are robust to controlling for soy production, which is a salient commodity during the study period both because of the surge in international prices, as well as the expansion of production in Brazil.

and farms as well as the average productivity in each sector are equilibrium outcomes that depend on the respective operating and entry costs. When entry costs in non-agriculture fall, average productivity in this sector increases as does the measure of firms in the market. This entry shifts the price of output as well as the relative price of capital to labor, inducing capital intensification in agriculture.

We calibrate our benchmark model to match key features of the Brazilian economy in the 1990s, at the beginning of our sample. We then run comparative static exercises where we lower the cost of entry in non-agriculture inducing new establishment entry, and study its implications for the organization of production in agriculture. Our model predicts that following a reduction in the non-agricultural entry cost—calibrated to generate a 1.4% increase in wages—agricultural employment falls by 1.57%, the number of farms declines by 1.38%, and capital per farm rises by over 5%. The large response in capital intensification is accompanied by land consolidation, with an increase in land-per farm of 1.4%.

These effects are economically meaningful. The increase in wages is consistent with the effects of large firms entries observed over the last three decades in Brazil. Our empirical strategy exploits rare episodes of large firm entries—defined as a new non-agricultural establishment comprising more than 0.01% of its micro-region's population in the year preceding the shock and with at least 200 employees at entry—and tracks their effects at the municipality level. Using a staggered difference-in-differences framework, we compare municipalities with a large entry ("epicenters") with municipalities in never-entered micro-regions. Since epicenters are more intensive in non-agriculture at baseline, we estimate effects on agriculture by comparing farms in surrounding municipalities to those in epicenter municipalities.[3]

We find that large non-agricultural firm entry causes formal non-agricultural employment to increase and raises wages in both agriculture and non-agriculture. Agricultural labor declines persistently, while the number of farms falls and land per farm increases. Consistent with capital-labor substitution, we also find increases in capital intensity, measured by tractor use, and a decline in the share of small farms. We document that these findings are primarily driven by entries in construction financed by the Federal Housing Program *Minha Casa Minha Vida*, the largest low-income housing construction and financing program in Brazil's history.[4]

To understand the contribution of the mechanization channel to these adjustments, we contrast the baseline results with a counterfactual in which the level of mechanization is held fixed. In this environment, the decline in agricultural employment is less than half as large, and capital per farm is unchanged. Farm exit still occurs, but labor intensity remains high, and production adjusts only through contraction. These results highlight that endogenous technology adoption is a critical mechanism through which "pull" shocks in non-agriculture generate reorganization and reallocation in agriculture. Importantly, capital deepening in agriculture emerges as both a response to and a mediator of labor market shocks, reinforcing

[3]See Section 2.4 for a full discussion. Effects of the large entry on surrounding municipalities relative to comparable municipalities in never-entered micro-regions are in progress.

[4]We are working on a complementary instrumental variables strategy that leverages Minha Casa Minha Vida as a exogenous source of increase in non-agricultural labor demand.

the structural transformation process.

Finally, using our quantitative model we plan to extend the analysis by exploring how initial differences in agricultural organization—such as landholding patterns and capital intensity—shape the response to non-agricultural entry. Brazil's regions vary widely in their reliance on labor- versus capital-intensive farming, as well as prevalence of smallholder versus large-scale farms, suggesting that identical shocks may generate heterogeneous adjustments. To capture this, we are working on recalibrating the model to reflect regional characteristics and simulate the same entry shock across these settings. This will allow us to test the model's predictions under different baseline conditions and shed light on how local production structures mediate the effects of labor market shocks on agricultural transformation.

Literature Review The debate about the role of technological progress in the agricultural and industrial sectors for the process of structural change dates back to the 1950s at least. Classical references like Harris and Todaro (1970) sustain that improvement in urban wages are fundamental to the shift in employment away from agriculture. At the same time, the proponents of the "food problem" as a central deterrent to sectorial reallocation, Schultz (1953), argue that agricultural productivity growth is key to releasing workers out of agriculture.

More recent work is also divided, Gollin et al. (2007) emphasizes agricultural productivity growth, while Boppart et al. (2023) emphasizes productivity growth elsewhere in the economy as driver of value-added and employment reallocation. Our study leverages micro-level evidence to argue that these two processes are linked together through endogenous technology adoption.

On the technical side, we propose a theory amenable to the discipline of modern micro-data. This discipline allows us to expense away from pure inference through relative prices, (Alvarez-Cuadrado and Poschke, 2011), while being consistent with aggregate patterns. Our analysis underscores the importance of relative input price changes influencing technology adoption in agriculture, building on insights from Manuelli and Seshadri (2014), Caunedo and Keller (2020) and Caunedo and Kala (2021).

Overall, a key implication of our analysis is that sluggish productivity growth in agriculture may not have its genesis in the agriculture sector. Hence, perhaps the "puzzle" to the sluggish adoption of technology in agriculture resides, to some extent, elsewhere in the economy Suri and Udry (2022).

2 The Impact of Non-agriculture Entry on the Agricultural Sector

2.1 Data

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 do informações So-

ciais (RAIS).[5] The unit of analysis are municipalities and the unit of time is a decade, since the Agricultural census runs every ten years.[6]

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, 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.

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 since 1986; the Agricultural Census for agricultural labor; and the Brazilian Population Census conducted in 1991, 2000, and 2010 to obtain information on informal employment as well as internal migration across municipalities.

From RAIS we aggregate total formal employment by sector at the municipality level. We also obtain information on the average years of schooling for workers in a given sector. Finally, using information on earnings, we compute average wage per hour and average monthly earnings in a given sector in units of the minimum wage. Wage per hour is computed from monthly earnings of a worker multiplied by 12, to obtain annual earnings, and divided by the total hours worked in a given year for each employee.

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.[7]

From the population Census we obtain information on the number of workers in agricultural and non-agricultural sectors, including informal ones. We further disaggregate those workers based on the level of their education into three groups – those with no or only (some) primary education, those with some secondary education, and those with some tertiary education. For this disaggregation we restrict the sample to individuals who are older than 10 years old and were working at some point of the year when the survey was conducted. Since both censuses are conducted every ten years, and our main outcomes stems

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

[6] While we refer to the unit of analysis as municipalities, the actual unit of analysis is Minimum Comparable Area (MCA) during study period, which are the smallest consistent sets of municipalities. 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. See https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/15771-evolucao-da-divisao-territorial-do-brasil.html. 1,016 of of the 5,507 observed in the Census and RAIS datasets in 1997-2017 are municipalities that split off from or joined other municipalities during this period.

[7] We use 1997 as the first year instead of 1996 as, prior to this year, not all municipalities were covered in RAIS.

from the Ag census, we match information from the 1991 population census with Ag census data in 1995-'97, the 2000 population census with Ag Census for 2006, and the 2010 population census with agricultural data for 2017.

The measure of informality follows the standard in the literature, Ulyssea (2018), i.e. the share of employed workers aged 18-64 with non-zero earnings that do not have a signed *Carteira de Trabalho* (employee registration). We use the 1991, 2000, and 2010 census to measure informality at the baseline entry decade as follows. Shocks between 2011 and 2016 have the 2010 as baseline, shocks between 2001 and 2010 have the 2000 census as baseline, and shocks before 2000 have 1991 as the baseline. Finally, we build measures of internal migration across municipalities from the population census, which we further disaggregated by individual socio-economic characteristics, like age or education level.

International Trade. We use data on imports and exports at the municipality-year-product level from Comex Stats, available for 1997-2022.[8] We aggregate these data at the municipality and year level, separately for agricultural and non-agricultural products.

Additional Data Sources We complement our analysis with micro-region-level exposure to import competition during Brazil's 1990-1994 trade liberalization, from Dix-Carneiro and Kovak (2017), which is used as an additional variable to balance control and treated units.

2.2 Motivating Facts

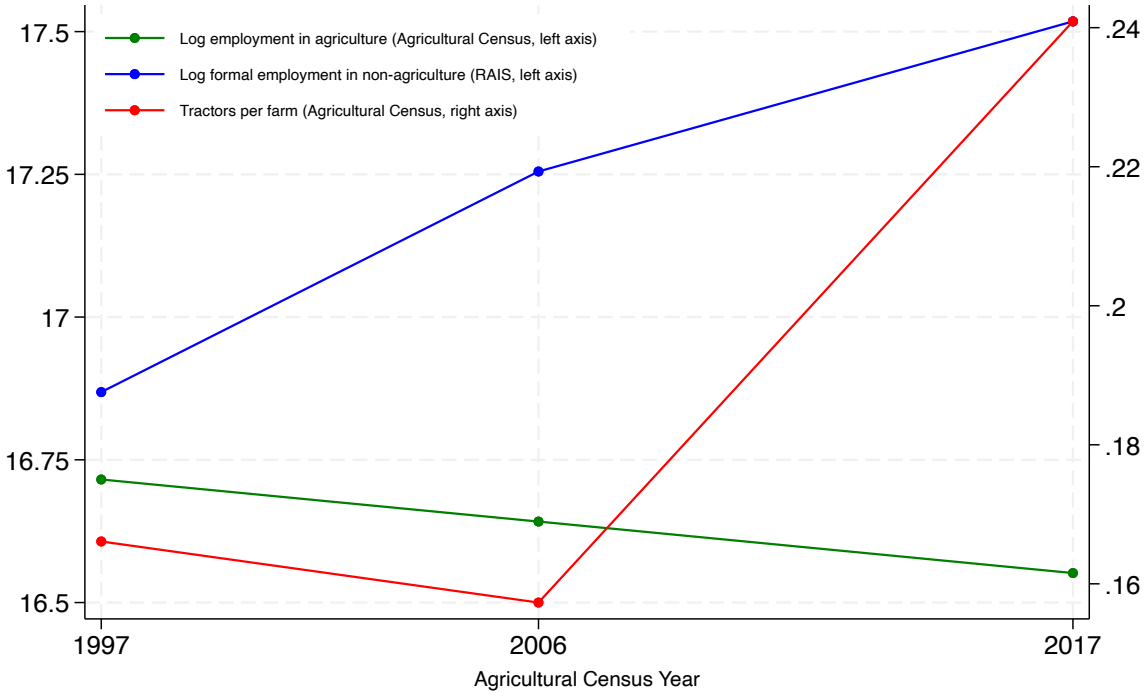
Before turning to our empirical strategy, we document a set of stylized facts that motivate our analysis and contextualize our identification approach. These facts, drawn from national and municipal-level data, illustrate the ongoing process of structural transformation in Brazil and provide suggestive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

Figure 1 shows standard patterns of structural transformation, with total employment in agriculture declining and (formal) employment in non-agriculture increasing, Herrendorf et al. (2014). At the same time, the number of tractors per agricultural worker rises steadily, reflecting increased mechanization in the agricultural sector. These patterns are consistent with a broader structural transformation, whereby labor reallocates away from agriculture while farms adopt more capital-intensive modes of production, Chen (2020).

At the municipal level, we can further investigate the proximate sources of such variation by bringing in data on wages. Higher agricultural wages are associated with greater mechanization, lower agricultural employment and higher formal non-agricultural employment. Figure 2 presents binned scatters plots using municipality-level data for each year in which the Ag Census was conducted. Panel A shows a positive relationship between agricultural wages and tractors per farm, and a negative relationship between

[8] Downloaded from https://basedosdados.org/dataset/br-me-comex-stat?bdm_table.

Figure 1: National trends in tractors per farm and employment



Notes: Data on tractors and agricultural employment are from the Agricultural Censuses for 1997, 2006, and 2017. Data for non-agricultural (formal) employment are from RAIS for the respective years.

agricultural wages and agricultural employment. Hence, municipalities where labor is relatively more expensive display more capital intensive farms and lower agricultural employment.

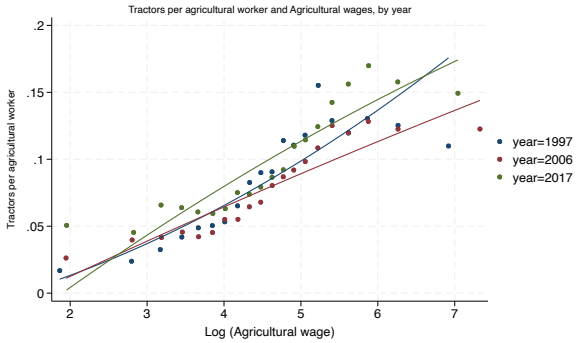
At the same time, Panel B of Figure 2 shows that municipalities with higher levels of formal non-agricultural employment also exhibit higher agricultural wages. Plausible explanation of observed patterns is that tighter labor markets in non-agriculture, or higher productivity in non-agriculture spills over into agriculture, putting upward pressure on rural wages. Hence, labor demand shocks originated 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.

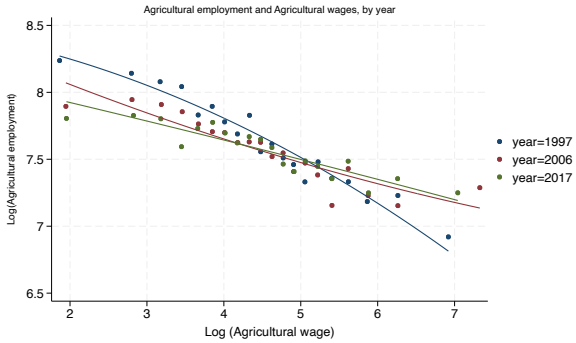
Since these are equilibrium outcomes in the labor and output markets, it is not possible to interpret these wage differentials, and associated outcomes, as causal. To learn more about the plausible causal role of wage increases on the organization of production in the agriculture sector, we study large firm entries

Figure 2: Organization of Agricultural Production and Local Labor Markets

Panel A: Municipal-level mechanization, wages, and employment in agriculture

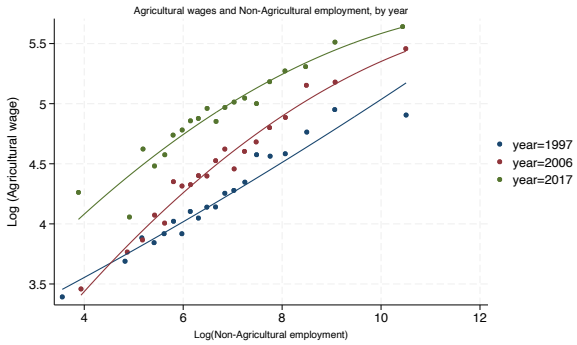


(a) Tractors per worker and ag wage

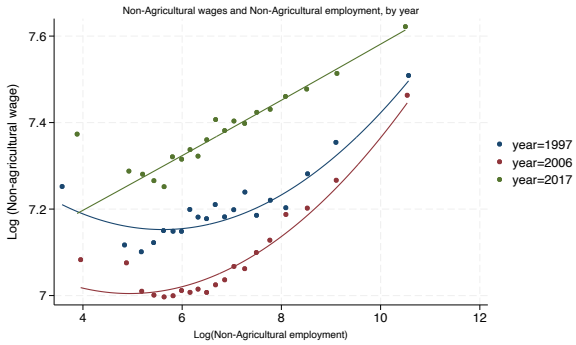


(b) Agricultural employment and ag wage

Panel B: Municipal-level wages and employment in non-agriculture



(c) Ag wage and non-ag employment



(d) Non-ag wage and non-ag employment

Notes: This figure plots binned scatters showing cross-municipality variation, separately by year and conditional on region fixed effects, for various municipal-level outcomes. Data on tractors, agricultural wages, and agricultural employment are from the Agricultural Censuses for 1997, 2006, and 2017. Data for non-agricultural wages and 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. Non-agricultural wages are measured as average worker monthly earnings and excluding workers in government. All wages are converted to 2017 reais using the iPCA inflation index.

into local labor markets and build a difference-in-differences research design.

2.3 Large-Firm Entries

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 0.01% of its micro-region's population in the year preceding the shock and with at least 200 employees at entry.[9] The former restriction identifies new entries that are large enough relative to a potential labor

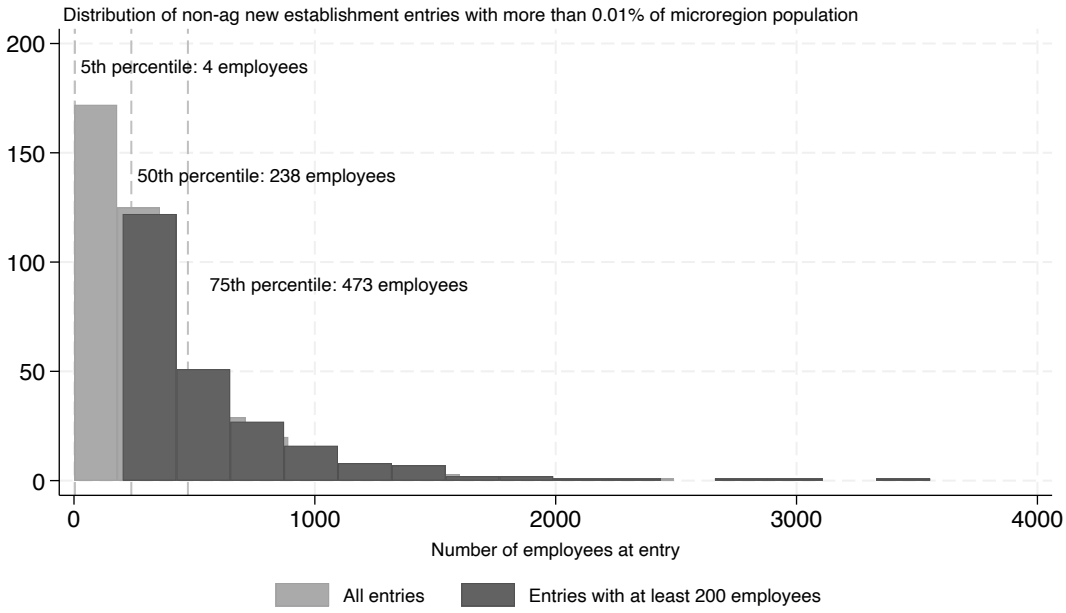
[9]This restriction is necessary to exclude new entries at very small micro-regions. Not doing so results in the inclusion of munici-

market of a region, whereas the second removes very small regions, for which even a 4-employee entry (5th percentile of new entering establishments) would otherwise be considered a large entry.

One 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.

Figure 3 displays the size distribution of all firm entries that account for at least 0.01% of the population in the micro-region, by the number of employees at entry. It also presents the distribution conditional on at least having 200 workers at entry. The median employment share accounted for by entry of new firms across municipalities is 238 employees, but the distribution is highly skewed with the 75th percentile of the distribution being 473 employees at entry, and many entry events with 1000+ employees. Our cutoff of 200 workers to define an entry shock is slightly below the median employment at entry.

Figure 3: Distribution of large entry shocks by number of employees and entry

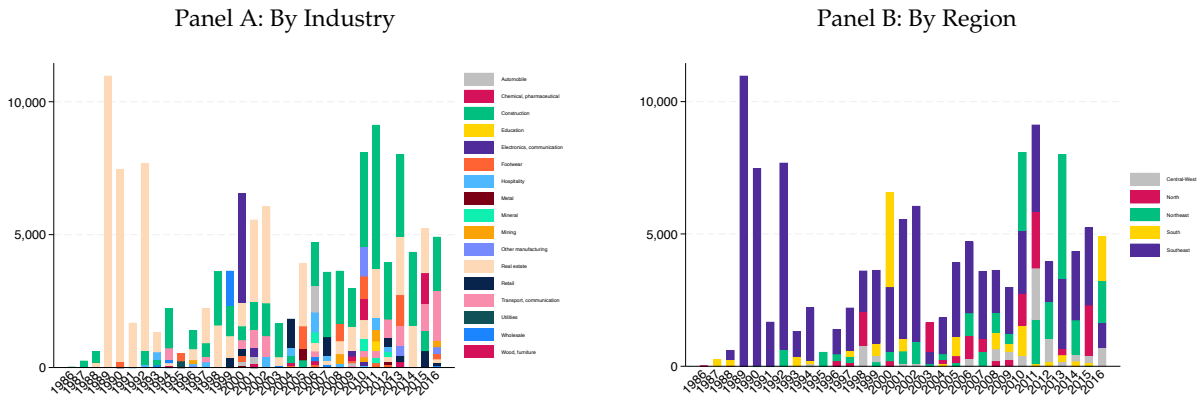


These large entries are unevenly distributed along industries and space. Figure 4 displays the distribution of large entry shocks through the years in our sample. Between the late '80s and the early 2000, most of these entries were concentrated in the real-state (in terms of employment contributions), followed by the construction sector and the transport and communication.

palities 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. We explore robustness to alternative thresholds, such as 100 or 300 employees.

Since the early 2000s, most of the employment contribution of these entry shocks stems from the construction sector. It is not surprising then, that the nature of our shocks is mostly concentrated in non-tradables, rather than in tradable industries, see Figure 4. In terms of spatial distribution, most of these large entries occurred in the Southeast, traditionally the richest modern area in Brazil. The Northeast, among the poorest regions in Brazil, is a close second, particularly since the 2000s.[10] This spatial heterogeneity, that is linked to economic development, provides an opportunity to study heterogeneous effects of firm entry.

Figure 4: Total Employment from Large Entries at Entry Year



Notes: A large entry is a new non-agricultural establishment with more than 0.1% of total micro-region population at entry and at least 200 workers. Excludes entries in manufacturing activities that directly demand agricultural inputs (i.e., food, beverages, textiles, and paper).

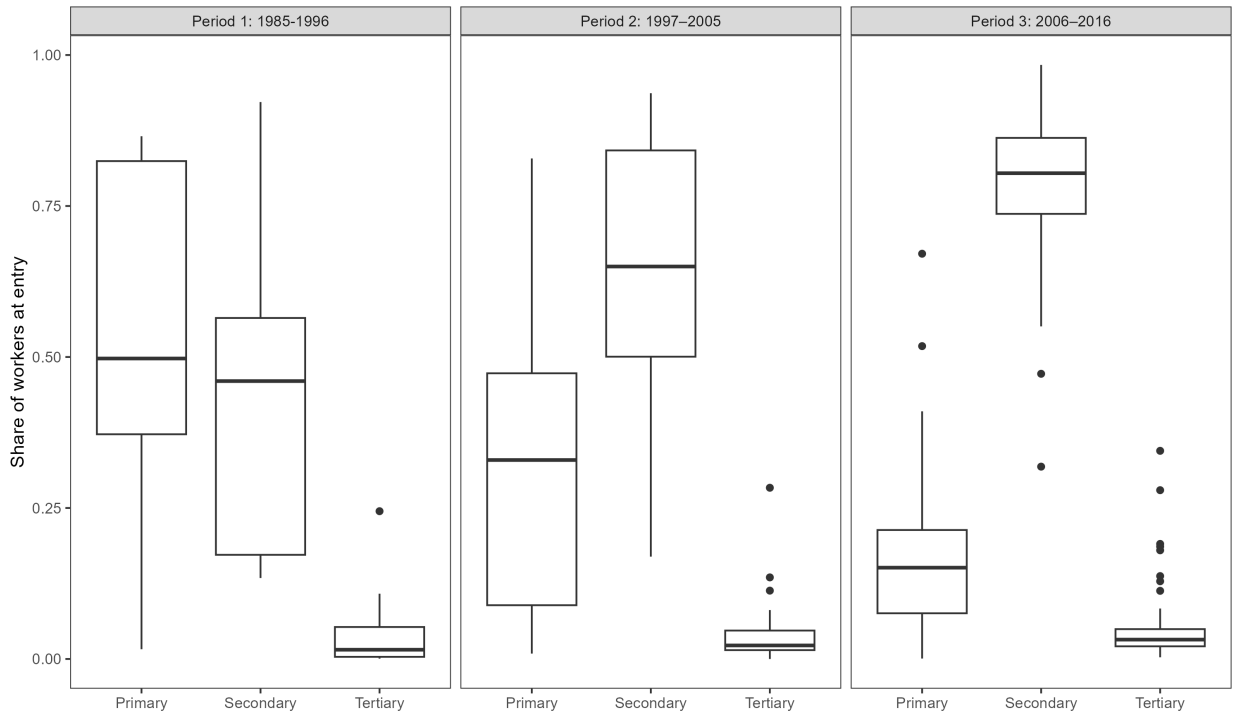
This spatial and sectoral distribution of shocks immediately draws two concerns. First, the Southeast–the region with high suitability for soy–in the early 2000s experienced a period of substantial expansion in this highly profitable crop, Bustos et al. (2016). Second, the expansion of the construction sector could be largely driven by higher agricultural profitability channeling into the real state market through better credit access, Bustos et al. (2020). To address these concerns, we run our main estimates separately for soy-producing and never-soy-producing municipalities, finding similar effects. We also use data from the Municipal Agricultural Survey (PAM), which provides monthly municipality-by-crop production and prices for the entire period, to document no pre-trends and no effects of large entries on soy-related production and prices.[11]

These shocks are also heterogeneous in the type of labor that these new firms demand. Figure 5 shows the workforce composition of these large entrants. The firms are overwhelmingly low- to medium-skill intensive, drawing primarily on workers with secondary education, and employing very few workers with tertiary education. This composition suggests that the effects of entry are likely to spill over into

[10] Appendix Figure A3 shows the distribution of entry shocks in space and through time.

[11] Future versions of this manuscript will include IV estimates from exogenous variation in construction entry through the Federal government's program Minha Casa Minha Vida, designed to address housing deficits in urban areas.

Figure 5: Workforce Composition of Large Firms entries



Notes: Skill distribution of a large firm entry shocks. A large entry is a new non-agricultural establishment with more than 0.1% of total micro-region population at entry and at least 200 workers. Skill categories refer to the maximum level of education attained. Primary education corresponds to middle school (8 years of formal education or less), secondary education corresponds to high school (11 years of formal education or less, but more than primary), and tertiary education corresponds to college (15 years of formal education in total).

nearby low-skill sectors, potentially including agriculture.

2.4 Empirical Strategy

As our primary interest is in the transformation of the agriculture sector in response to a large entry shock in non-agriculture, we combine RAIS data on formal entry in non-agriculture with information from the Ag Census. The Ag Census provides a richer set of outcomes and includes all agricultural establishments across the country, including informal ones, as well as measures of mechanization and farm characteristics, i.e. land size, crops, etc.

Design. We start by defining municipalities that ever experienced a large entry shock between 1985 and 2016 as an *epicenter*, municipalities that never experienced a large entry but are in the same micro-region as the *surroundings*, and municipalities that never experienced a large entry but are in different micro-regions as *never-entered regions*.

This distinction is useful because epicenters tend to be more non-agriculture intensive than surroundings at baseline, yet they are part of a largely integrated local labor market (by transportation systems,

etc.), the micro-region. Figure 6 characterize the share of the population in each municipality working in the agriculture sector. The average agricultural employment share in the surroundings is 37% while in the epicenter is only 12%. This is consistent with epicenter municipalities being substantially more urban than their surrounding counterparts, see Figure A4. As a result, we can expect much of the effect of large entries on tractor adoption to happen in agriculture-intensive surrounding farms instead of in already tractor-adopting farms in the epicenters.

To estimate the effect of large entries on non-agriculture outcomes, we compare *epicenters* relative to municipalities in *never-entered regions*. A strength of this comparison is that the control group is never affected by large entries, and the high-frequency of the data on non-agricultural outcomes (annual for 1985-2017) allows us to implement a Staggered Differences-in-Differences with heterogeneous treatment effects a la De Chaisemartin and d'Haultfoeuille (2020). Its corresponding TWFE estimator, easier to interpret, is:

$$y_{mt} = \sum_{k=-5, k\neq 0}^{k=10} \beta_k \left[\text{Epicenter}_m \times (1_{k=t-t_m^*+1})\right] + \delta_t + \delta_m + \epsilon_{mt}$$

where m is a municipality, Epicenter is a dummy with value 1 if the municipality ever experienced a large entry, zero if the municipality is in a never-entered region, t_m^* denotes the years of the shock, i.e. the first year where the firm is recorded in the data.

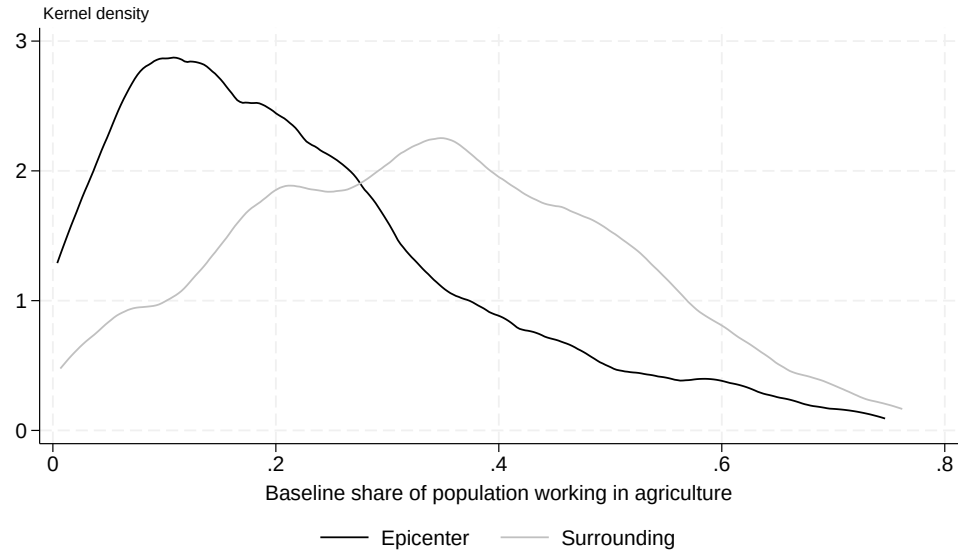
To estimate the effect on agricultural outcomes we compare the farms in *surroundings* to those, already more mechanized, in the *epicenters*. On the one hand, an attractive feature of this comparison is that it treats the micro-region as the relevant local labor market, focusing on within-market cross-municipality reallocation forces towards the epicenter, into which labor is pulled. On the other hand, its weakness is that farms in the epicenter are also treated, and as a result our effects on agricultural outcomes might be underestimates of the true effects. We are aware of this limitation and are working on the cleaner design that compares agricultural outcomes at either epicenters or surroundings with comparable municipalities in never-entered regions.[12]

In addition, because the Agricultural Census is conducted only once every ten years, our analysis of agricultural outcomes must be based on a decadal frequency. This coarse timing poses a challenge for isolating the timing and sequencing of shocks: within each inter-census decade, multiple entry events cannot be separately timed and are instead aggregated into a single measure for the decade preceding the Census year. In practice, 79% of the micro-regions in our sample have only one epicenter in the relevant decade.[13] To accomodate the minority of cases where a micro-region contains multiple epicenters, we

[12] A challenge in mirroring this design for agricultural outcomes is that agriculture-intensive municipalities are very heterogeneous, making it hard to find appropriate control municipalities in never-entered regions when we only have three years for the agricultural census (1997, 2006, 2017),. In contrast, data for non-agricultural outcomes are available at the annual level, facilitating a staggered design that compares epicenters to never-treated regions.

[13] All micro-regions have at most four epicenters. In addition, 98% of municipalities (119 out of 122) with any large entry shock had these shocks occurring within the same calendar year. If a municipality had large entry shocks in multiple years, we consider the first entry shock as the original treatment, and report all results for that epicenter as relative to the year of that first entry shock. Within any given year, 90% of municipalities with any large entry shock experience one shock, and there are no municipalities with more than two large entry shocks in any given year.

Figure 6: Baseline employment in agriculture, epicenter vs. surroundings



Notes: The baseline corresponds to the year prior to the large entry shock.

examine heterogeneity by the size of the employment shock, defined as total employment at entry summed across all (shock) entries in the preceding decade.

Given the lower frequency of data on agricultural outcomes, we estimate the effect of large entries on these outcomes using the standard TWFE staggared difference-in-differences regression, which imposes homogenous treatment effects across entry cohorts.[14] Specifically, we estimate:

$$y_{mt} = \sum_{d=-2, d\neq -1}^{2} \beta_d \left[\text{Surrounding}_m \times (t_m^* \in \Theta_d)\right] + \delta_d + \delta_t + \delta_m + \epsilon_{mt}$$

where again, m denotes the municipality, t_m^* is the time of the shock, Surrounding is a dummy with value 1 if the municipality is in the surrounding of a shock and zero if the municipality is an epicenter municipality, d corresponds to the decade relative to the shock and Θ_d is the years in the corresponding decade d. For the 1997-2006 entry cohort, $k = 0$ in 1997, $k = 1$ in 2006, and $k = 2$ in 2017. Hence, for the 2006-2017 entry cohort, $k = -1$ in 1997, $k = 0$ in 2006, and $k = 1$ in 2017.

For all specifications, wee cluster standard errors at the micro-region level.

2.5 Results

We are now ready to describe the effect of entry shocks into the economy. We start by describing the magnitude of the shock and then move onto its impact on wages, allocations of workers across sectors,

[14] We report robustness to the De Chaisemartin and d'Haultfoeuille (2020) estimator for key outcomes such as wages and tractors.

and finally organization of production in agriculture.

The staggered difference-in-difference approach allows us to interpret these effects as the causal impact of entry on outcomes—relative to never-entered regions for non-agricultural outcomes and relative to never-entered surrounding municipalities in the same micro-region for agricultural outcomes. We present results at an annual level whenever possible, and at a decade level otherwise.

Figure 7 summarizes the main effects of the shock on the labor market. Figure 7a summarizes the magnitude of the non-agriculture employment effect associated to the large entry shock, inducing 549 additional FTE workers relative to the surroundings. In total, the entry shock induces an increase in employment of 15.1% in the epicenter relative to the surroundings, Figure 7b.

This employment effect in the epicenter is accompanied by structural transformation. Figure 7c shows that most of the increase is accounted for employment gains in non-agriculture (18%), while the agriculture sector looses employment in the epicenter relative to its surroundings (4% on impact and 15% over a decade), Figure 7d. Since these are relative effects, our finding suggests that even if both epicenter and surrounding are loosing employment in agriculture due to structural transformation, the loses in the epicenter are stronger than in the surroundings.

Accompanying the increase in employment are higher wages in non-agriculture in the epicenter relative to the surroundings (6p.p. higher differential than at baseline). Wages in agriculture fall in the epicenter relative to the surroundings (1.5% on impact), consistently possibly with higher wages in the surroundings. However, the point estimates are noisy. We return to this feature when analyzing data from the Ag Census, which also includes informal agriculture.

Since the demand shock mostly affects non-agriculture workers and those shocks are concentrated in the real state and construction sectors, it is useful to understand the nature of the demand shift in term of workers' human capital. We proxy the latter by educational attainment of the worker. Figures 8a, 8c and 8e show that there is widespread increases in real wages across all educational groups, and the highest improvement is for workers with some primary education, 10p.p., versus 7p.p. and 7.6p.p. for workers with secondary and tertiary education, respectively. This increase in wages lasts at most three years, consistent with the wining down of the non-agriculture employment effects documented earlier.

Despite the large entry shock being mostly accounted for by establishment intensive in workers with secondary effects, the largest employment effects are experienced by workers with primary education.

We know that construction and real state entries drive the bulk of the new employment from entry. When we study the impact of the entry shock on employment by industry, we find that the bulk of the employment effect on low-skill workers is in the Construction sector, see Figure A5d. More broadly, Figures A5a and A5b show that most of the effects stem from non-tradable sectors rather than tradable sector, where multinational entry is prevalent. In addition, despite the real state sector accounting for a large share of entry, particularly at the beginning of the sample period, we found no substantial employment effects of these types of entry for any educational group.

Having characterized the effect of these shocks on the formal labor market, we can now explore impacts in agriculture more broadly, including land allocations and capital intensification. To do so, we move to our second design and explore outcomes across Ag Census decades.

Table 1 presents the effect of large entry shocks on agricultural wages, agricultural employment and tractors per farm for the surroundings relative to the epicenter, allowing for heterogeneity in the size of the entry shock in terms of employment and for heterogeneity by tractor intensity (measured as tractors per farm) in surrounding municipalities the decade prior to entry. We find that the entry shock increases real agricultural wages in the surroundings by 0.30 log points, while decreasing agricultural labor by 0.30 log points. These results are consistent with the effects documented on annual employment looking at formal agricultural employment (with less noise in the estimates for employment).

Table 1: Effect of large entry shock on agricultural outcomes in surroundings, with heterogeneity by baseline tractor intensity and log entry shock employment

	Log wage	Log workers	Log tractors/farm	Log # small farms
Sur X Post	0.303**	-0.307**	0.0544*	-0.351**
	(0.136)	(0.146)	(0.0279)	(0.144)
Sur X Post X Base Tractors/farm	-0.0512	0.138**	0.101***	0.00634
	(0.0951)	(0.0683)	(0.0356)	(0.159)
Sur X Post X Entry Size	-0.0000586	0.0000475	-0.0000358**	0.000111***
	(0.0000477)	(0.0000291)	(0.0000144)	(0.0000412)
Observations	2087	2092	2087	2092

Lower labor and increasing wages induces higher mechanization, with an increase of 0.05 log points in tractors per farm. Consistent with endogenous mechanization, we find that the effects on wages are driven entirely by surrounding farms with little or no tractor intensity at baseline. Surprisingly, tractor intensity at baseline attenuates the negative employment effects and amplifies tractor adoption. This could be explained by baseline differences in soil ruggedness / appropriateness for tractors, which we don't control for but will be able to in the future. It also appears that all effects are slightly attenuated in regions where the large the entry shock is most sizable. This could be driven by demand effects towards agriculture in the epicenter, which would mute the relative effect documented for the surrounding.

We also explore the effects on this key outcomes in the agriculture sector across different decades after the entry shock exploiting a two-way fixed effect design. Figure 9 shows the magnitude of the effects on real wages in agriculture, labor and tractors per farm. The effects both on real wages as well as on tractors per farm persist two decades after the shock. The effects on agricultural labor are noisier but the point estimates suggest loss in employment in the surroundings relative to the epicenter.

Mechanization is a function not only of relative prices, but also of production scales, Foster and Rosenzweig (2022). It is expected then that the same shock to the price of labor in agriculture would generate heterogeneous effects on mechanization depending on production scales. Table 2 present results when we

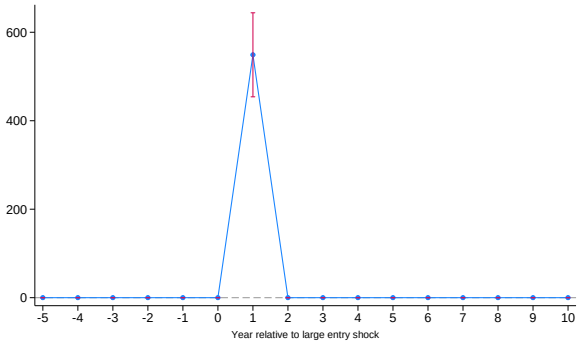
Table 2: Effect of large entry shock, with heterogeneity by surroundings' baseline average farm size

	Log ag wage	Log ag labor	Log tractors per farm
SpillXpost	0.151	-0.237*	0.00545
	(0.0966)	(0.137)	(0.0144)
SpillXpostXBLFshare	0.570*	0.902**	0.308**
	(0.322)	(0.411)	(0.152)
Observations	2087	2092	2087

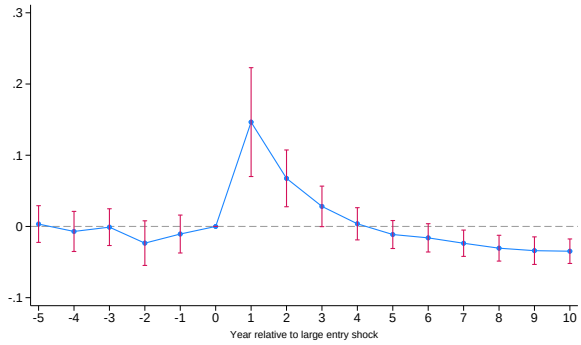
interact the large entry shock with average land holdings in the municipality. We find two key patterns. First, municipalities where farm holdings are relatively large at baseline are the ones that mechanize the most, consistently with scale being important for the elasticity of mechanization to the shock. Second, these municipalities are also those where increases in agricultural wages concentrate. Interestingly, while the direct effect of the shock on agricultural labor in the surroundings is negative, the shock generates relatively higher labor demand in the surroundings to the epicenter where land-holdings are larger. One explanation for this finding is that where land-holdings are larger, most of the labor demand effect reallocates workers from agriculture in the epicenter, rather than the surroundings.

Figure 7: Impact of Large Firm Entry Shock

Panel A: Shock

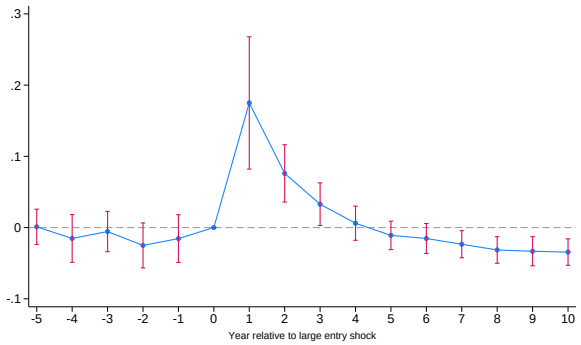


(a) Non-agriculture employment per establishment.

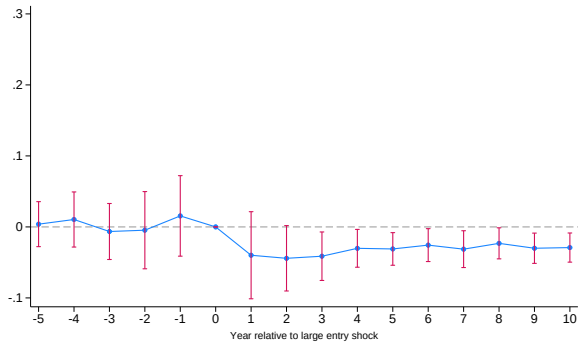


(b) Total Employment (log).

Panel B: Sectorial Employment

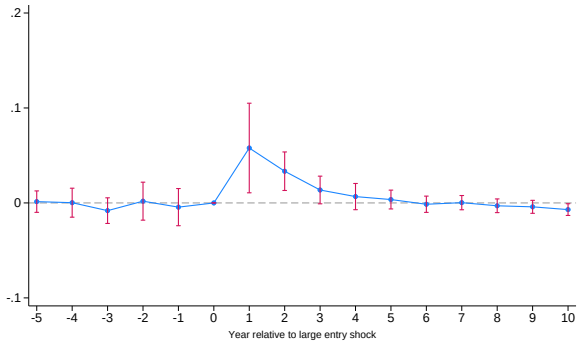


(c) Non-agriculture employment (log).

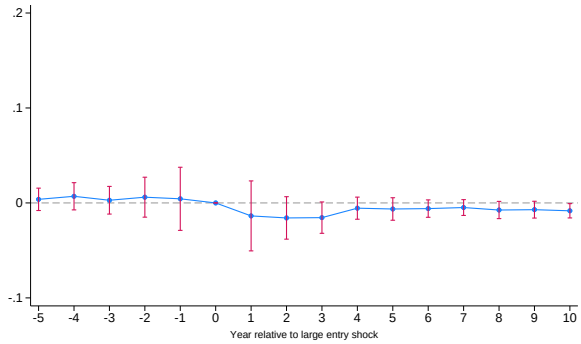


(d) Agriculture employment (log).

Panel C: Real Wages (average)



(e) Wages in non-agriculture (log)

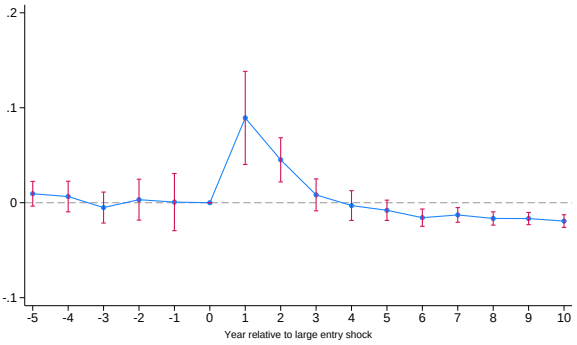


(f) Wages in agriculture (log)

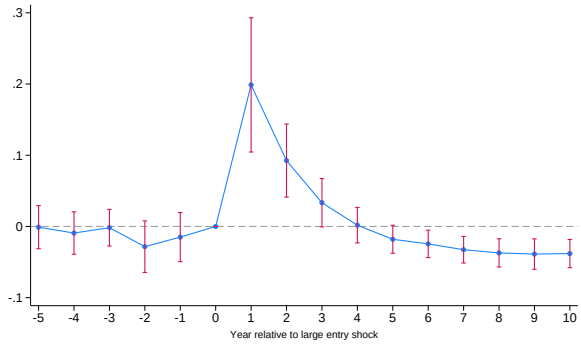
Notes: This figure presents difference-in-differences estimates of a large entry shock on epicenter municipality outcomes relative to surrounding municipalities. Estimates allow for heterogeneous treatment effects, use municipalities in never-entered as controls, a la De Chaisemartin and d'Haultfoeuille (2020), and normalized coefficients, such that the point estimate at k is the effect of an increase in one unit of treatment at k. A large entry is a new non-agricultural establishment with more than 0.01% of total micro-region population and at least 200 employees at entry. Excludes entries in manufacturing activities that directly demand agricultural inputs (i.e., food, beverages, textiles, and paper). Sample includes only formal sector outcomes. Real wages are reported in constant Brazilian Reais as of August 2017.

Figure 8: Large Firm Entry Shock on Wage and Employment by Education Group

Panel A: Primary education or less

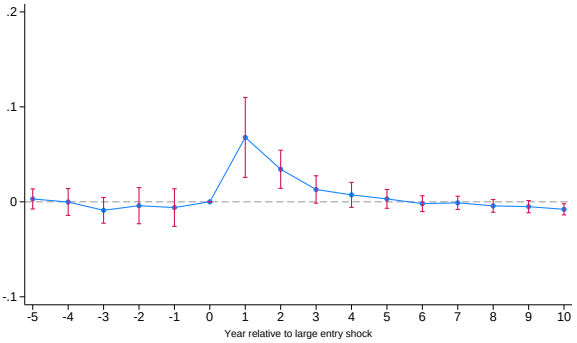


(a) Average wage (log)

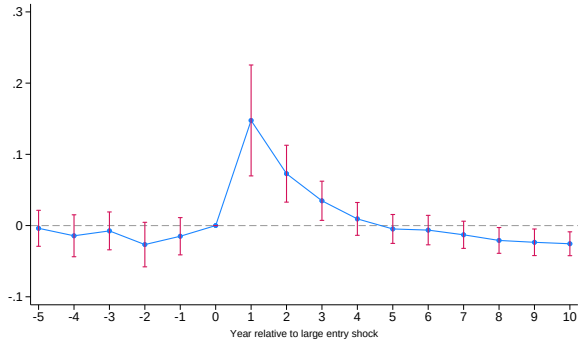


(b) Employment (log)

Panel B: Secondary Education or Less

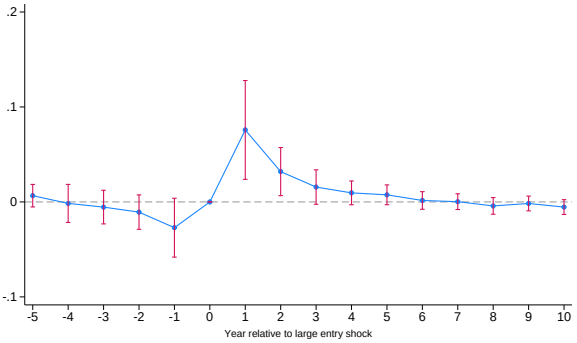


(c) Average wage (log)

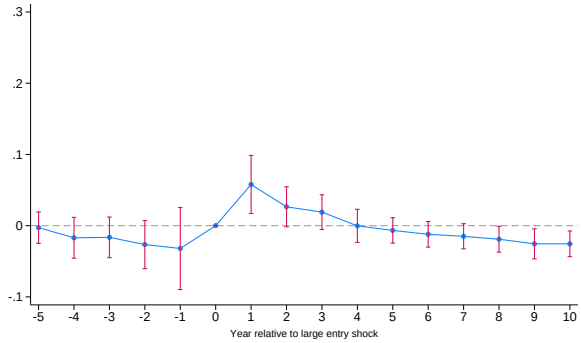


(d) Employment (log)

Panel C: Tertiary education of less



(e) Average wage (log)

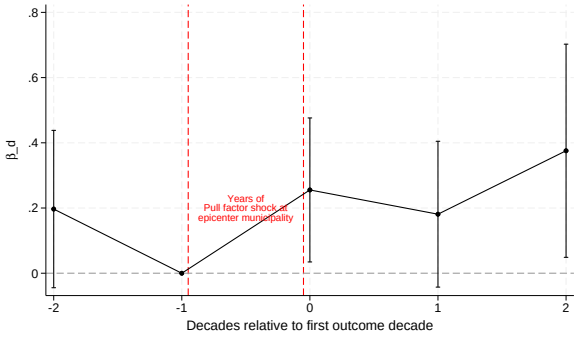


(f) Employment (log)

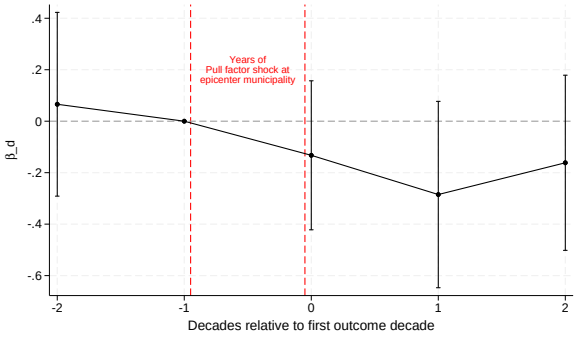
Notes:This figure presents difference-in-differences estimates of a large entry shock on epicenter municipality outcomes relative to surrounding municipalities. Estimates allow for heterogeneous treatment effects and use never-treated surrounding municipalities as controls, a la De Chaisemartin and d'Haultfoeuille (2020). A large entry is a new non-agricultural establishment with more than 0.1% of total micro-region population and at least 200 employees at entry. Excludes entries in manufacturing activities that directly demand agricultural inputs (i.e., food, beverages, textiles, and paper). Real wages are reported in constant Brazilian Reais as of August 2017.

Figure 9: Key outcomes in agriculture, two-way fixed effects.

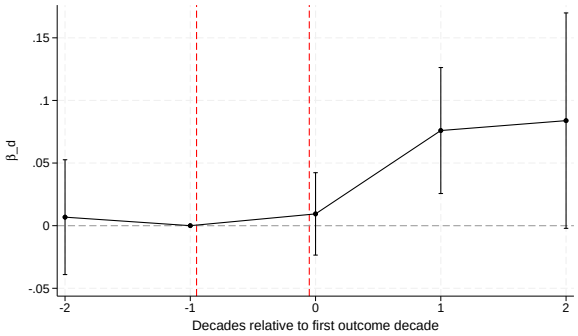
Panel A: Real wages agriculture



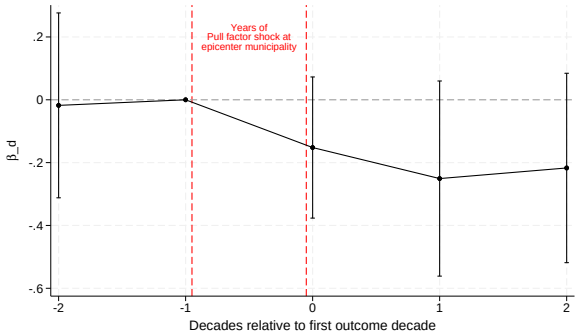
Panel B: Agricultural Labor



Panel C: Tractors per farm



Panel D: Small farms



Notes: Results from two-way fixed effects, surroundings vs. epicenter, in response to a large entry shock.

3 A Model of Firms, Farms and Input Intensification

This is a parsimonious two sector model of heterogeneous firms in the non-agriculture sector, and farms in the agriculture sector. Distinctively, capital-intensity in the agriculture sector is endogenously chosen and responds to shifts in relative prices. The model aggregates to a canonical two sector economy on the production side, and displays non-homotheticities in preferences. We make a small departure from the canonical model to include an informal sector in non-agriculture, where a representative firm operates. This additional sector allows us to match particular features of the Brazilian economy but is otherwise inconsequential to the main theoretical results.

3.1 Preferences and Technology

Preferences. Consider a complete markets economy populated by a measure $\bar{N}$ of households that consume and supply labor inelastically. Preferences are constant elasticity aggregators of agriculture and non-agriculture goods. Non-agriculture goods are produced in the formal and the informal sector and are differentiated products from the point of view of the consumer.

The problem of the representative household is

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

subject to

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

$$b_0 \geq 0 \quad \text{given},$$

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. There are three sectors of production, agriculture (a), non-agriculture (n), and a non-agriculture informal sector (s). There is a mass M_n of heterogeneous firms that produce a homogeneous good with a decreasing returns to scale technology in labor. Since firms' problems are static, we drop the time subscript in their description.

Firm's i output is

$$Y_{n_i} = Z_{n_i} N_{n_i}^{\alpha},$$

where Z_{n_i} is a Hicks Neutral productivity term, and α is a measure of decreasing returns to scale.

In the informal sector there is a representative firm that produces with a decreasing returns to scale technology in labor

$$Y_s = Z_s N_s^{\alpha},$$

where Z_s is a Hicks Neutral productivity term, and α is the degree of decreasing returns to scale, equal to that in formal non-agriculture, for simplicity.[15]

In the agricultural sector there is a mass M_a of heterogeneous farms operating a decreasing returns to scale technology that uses land, capital and labor as inputs. Farming requires performing a set of activities b of measure 1, each with output $x(b)$, and total output is

$$Y_{a_j} = Z_{a_j} \left(e^{\int_{b=0}^{1} \ln x_j(b) db} \right)^{\gamma} L_j^{\gamma_L}.$$

where Z_{a_j} is a Hicks Neutral productivity term, L_j is the land allocated to farm j. The output elasticity to land is γ_L. Activities can be performed with labor or capital $x_j(b) = a_k(b)k_j(b) + a_n(b)n_j(b)$, capturing the different suitability of labor and capital in completing them. This suitability is summarized through a profile of comparative advantage that tracks the marginal product of a unit of capital relative to labor in completing a task, $\frac{a_k(b)}{a_n(b)}$.

Assumption 1 $\frac{a_n(b)}{a_k(b)}$ is continuously differentiable and decreasing in b.

Hence, tasks with lower index b are less likely to be mechanized, because those tasks are those where labor is the most productive.

Operation costs and Free entry. Each firm in non-agriculture and each agricultural farm pays fixed costs in units of labor, f_n and f_a, respectively. Firms and farms, can enter the market after paying a fixed cost in units of labor, F_n and F_a, respectively; and upon entry, observe their productivity draw Z_{o_i} from a known distribution $G_o(Z)$ for $o \in \{a, n\}$. There is free entry in each sector.

Feasibility. For each period of time t the following feasibility constraints are satisfied. Labor can be allocated to production, N^p, or to pay for operating and entry costs, N^f.

$$N^p \equiv M_n \int_0^{\infty} N_n(Z) g_n(Z) dZ + N_s + M_a \int_0^{\infty} N_a(Z) g_a(Z) dZ,$$

$$N^f \equiv F_a m_a + F_n m_n + f_a M_a + f_n M_n.$$

The labor feasibility constraint is $N^p + N^f = \bar{N}$.

The feasibility constraint for capital and land are simply,

$$\int_{j \in M_a} K_j = K \qquad \text{and} \qquad \int_{j \in M_a} L_j = L \equiv \bar{N} l.$$

[15] For the rest of this section, we use a firm in non-agriculture as a reference to *formal* non-agricultural sector, n, unless we explicitly specify.

Finally, the feasibility constraints for goods are

$$C_a = \int_{j \in M_a} Y_{a_j} \equiv Y_a, \qquad C_s = Y_s, \qquad C_n + X = \int_{i \in M_n} Y_{n_i} \equiv Y_n.$$

The law of motion for capital and firms satisfy

$$\bar{N} b_{t+1} = X_t,$$

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

$$M_{at+1} = m_t^a + M_{at}(1 - \delta_{M_a}),$$

$$M_{nt+1} = m_t^n + M_{nt}(1 - \delta_{M_n}),$$

where $\delta_{M_a}, \delta_{M_n}$ are the exogenous exit rates for farms and firms, respectively.

3.2 Equilibrium Allocations

Mechanization choices. Since output from tasks performed either with labor or capital is additive, there is full specialization, i.e. some tasks are performed entirely with labor, some entirely with capital.

Activities are performed with labor, whenever

$$\frac{W_t}{a_n(b)} \leq \frac{R_{kt}}{a_k(b)}.$$

Hence, given Assumption 1, there exist a unique threshold β_t, such that if $b < \beta_t$ such an activity is performed with labor and otherwise, it is performed with capital. An important feature of this technology is that despite heterogeneity in Hicks neutral productivity (and therefore in optimal land-sizes), the choice of mechanization depends solely on market prices and the profile of comparative advantage across tasks, common across farms as in Caunedo and Kala (2021). In other words, the share of tasks that are mechanized is the same across farms.[16]

Replacing back into the production technology,

$$Y_{a_j} = Z_{a_j} \mathcal{A}(\beta_j) \left(K_j^{1-\beta_t} N_{a_j}^{\beta_t} \right)^{\gamma} L_j^{\gamma_L}, \qquad \text{for } \gamma + \gamma_L < 1,$$

for $N_{a_j} \equiv \int_{b=0}^{\beta_t} n_{a_j}(b)$, $K_j \equiv \int_{b=\beta_t}^{1} k_{a_j}(b)$ and the threshold implicitly defined by $\frac{W_t}{a_n(\beta_t)} = \frac{R_t}{a_k(\beta_t)}$. The output elasticity to labor is $\beta_j \gamma$, and to capital is $(1 - \beta_j)\gamma$. Farm productivity depends endogenously on technology choices through $\mathcal{A}(\beta_j) \equiv \left(\int_{b=0}^{\beta_t} a_n(b) \right)^{\beta_j} \left(\int_{b=\beta_t}^{1} a_k(b) \right)^{1-\beta_j}$. When the price of capital and labor is the same across farms, $\beta_j = \beta$, independent of other farm characteristics.

[16] This feature brings substantial tractability to our problem and allows us to characterize aggregates.

Firm and farm survival. The presence of operating costs and decreasing returns to scale in determines threshold productivities, $\underline{Z}_{nt}$, and $\underline{Z}_{at}$ for the marginal firm and farm, respectively.

$$\tilde{\pi}_{nt}(\underline{Z}_{nt}) = W_t f_n, \qquad \tilde{\pi}_{at}(\underline{Z}_{at}) = W_t f_a$$

where $\tilde{\pi}$ are the variable component of firm profits, $\pi_{ot} = \tilde{\pi}_{ot} - W_t f_o = (1-\alpha) P_{ot} Y_{ot} - W_t f_o$.

Defining the distribution productivity conditional on survival as $g_o(Z) = \frac{\tilde{g}_o(Z)}{1-G_o(\underline{Z})}$ for $\tilde{g}$ the p.d.f. associated with G in each sector $o \in \{n, a\}$. Therefore, $g_o \geq 0$ if $Z > \underline{Z}_o$ and zero otherwise. In addition, we define the productivity of the "average" firm as $Z_n \equiv \left(\int_0^\infty Z_{n_i}^{\frac{1}{1-\alpha}} g_n(Z_{n_i}) dZ_{n_i} \right)^{1-\alpha}$. This average productivity is independent of the measure of firms in the market and depends on the conditional probability of each productivity level, g_n. We use an analogous expression in the agriculture sector, $Z_a \equiv \left(\int_0^\infty Z_{a_j}^{\frac{1}{1-\alpha}} g_a(Z_{a_j}) dZ_{a_j} \right)^{1-\alpha}$.

Employment. The optimality conditions for employment allocation imply that relative productivity across firms and farms track relative employment allocations

$$\frac{Z_{n_i t}}{Z_{n_{i'} t}} = \left(\frac{N_{n_i t}}{N_{n_{i'} t}} \right)^{1-\alpha}, \qquad \text{and} \qquad \frac{Z_{a_j t}}{Z_{a_{j'} t}} = \left(\frac{N_{a_j}}{N_{a_{j'}}} \right)^{1-\gamma-\gamma_L}.$$

Relative employment allocations in sectors s and e are similar to the ones above, but also a function of equilibrium relative prices,

$$\frac{Z_{st}}{Z_{nt}} = \left(\frac{N_{st}}{N_{nt}} \right)^{1-\alpha} \frac{P_{nt}}{P_{st}}.$$

Finally, we can characterize the relative employment allocation between agriculture and informal non-agriculture as a function of labor productivities,

$$\frac{Y_{at}/N_{at}}{Y_{st}/N_{st}} = \frac{\alpha}{\gamma\beta} \frac{P_{st}}{P_{at}},$$

where Y_{at}/N_{at} is the labor productivity of the average farm in the agriculture sector.

Prices. The optimality conditions for household consumption and associated feasibility constraints yield expressions for sectorial prices, $P_{lt} = \frac{\theta_l}{C_{lt} - \bar{c}_l}$ for $l \in \{a, n, s\}$ with

$$P_{nt} = \frac{\theta_n}{M_{nt} Z_{nt} N_{nt}^\alpha - X_t - \bar{c}_n} = \frac{\theta_n}{M_{nt} Z_{nt} N_{nt}^\alpha (1 - s_t) - \bar{c}_n}.$$

We define s as the savings rate in the economy, i.e. the share of output from sector n devoted to investment rather than consumption.

The price in agriculture is

$$P_{at} = \frac{\theta_a}{M_{at} Z_{at} \mathcal{A}(\beta_j)(N_{at}^{\beta_t} K_t^{1-\beta_t})^{\gamma} L_t^{\gamma_L} - \bar{c}_a}. \tag{1}$$

Replacing these prices into the optimality condition for labor yields

$$\frac{Z_s}{Z_n} = \left(\frac{N_s}{N_n}\right)^{1-\alpha} \frac{\theta_n}{\theta_s} \frac{Z_s N_s^{\alpha} - \bar{c}_s}{(1-s) M_n Z_n N_n^{\alpha} - \bar{c}_n}.$$

An implication of the optimal employment allocation in the non-agricultural sector, is that whenever entry in non-agriculture induces higher employment for the average firm, it also increases employment in the informal sector. Absent non-homotheticities, employment levels are independent of average productivity.

The amount of labor allocated to productive activities is in turn endogenous and depends on equilibrium entry and the measure of firms in the market, through the labor market clearing condition:

$$\bar{N}_t^p + F_a m_{at} + F_n m_{nt} + f_a M_{at} + f_n M_{nt} = \bar{N}.$$

Free Entry. Entry in each market, m, occurs until the entry fixed cost equals the expected value of a firm in the market, v, conditional on a survival. A firm operates if the initial productivity draw is above the threshold productivity, and the firm is not hit by the exogenous exit probability.

$$W_t F_n = (1 - G_n(\underline{Z}_{nt+1}))\nu \int_{\underline{Z}_{nt+1}}^{\infty} v_{nt+1}(Z) g_{nt+1}(Z) dZ,$$

$$W_t F_a = (1 - G_a(\underline{Z}_{at+1}))\nu \int_{\underline{Z}_a}^{\infty} v_{at+1}(Z) g_{at+1}(Z) dZ.$$

The expected value of a firm is simply the discounted expected value of profits, $v_{ot}(Z) = \sum_t^{\infty} \nu^t (1 - \delta_{M_o})^t \pi_{ot}(Z)$ where the discount factor is the hazard rate of firms in that sector as well as the household discount factor (who own the firms in the market).

Zero profit condition. The zero-profit condition in turn determines the productivity of the marginal firm and the marginal farm.

$$(1-\alpha) Z_{nt} N_{nt}^{\alpha} \left(\frac{\underline{Z}_{nt}}{Z_{nt}}\right)^{\frac{1}{1-\alpha}} = W_t f_n. \tag{2}$$

$$(1 - \gamma - \gamma_L) P_{at} K_t^{(1-\beta_t)\gamma} N_{at}^{\beta\gamma_t} L_t^{\gamma_L} Z_{at} \mathcal{A}(\beta_j) \left(\frac{\underline{Z}_{at}}{Z_{at}}\right)^{\frac{1}{1-\gamma-\gamma_L}} = W_t f_a. \tag{3}$$

Capital accumulation. We normalize the price of non-agriculture goods to one, hence, all dynamic decisions are expressed in units of the marginal utility of consumption for non-agriculture goods, $P_{nt} = \frac{\theta_n}{C_{nt} - \bar{c}_n}$. The optimal capital allocation follows from the Euler equation of the households, $P_{nt} = \beta R_{bt+1} P_{nt+1}$.

The optimal accumulation of capital follows

$$\lambda_{nt} = \beta \left(R_{kt+1} + (1 - \delta_k) \lambda_{nt+1} \right)$$

and the return to capital is the marginal product of capital in the agriculture sector.

3.3 Stationary Equilibrium

In the steady state, prices and allocations are constant, so β is constant, and we can normalize the profile of comparative advantage so that $\mathcal{A}(\beta) = 1$.

The measure of farms and firms are constant, and the entry rate is pin down from the law of motion of firms and farms

$$m_a = \delta_{M_a} M_a, \qquad m_n = \delta_{M_n} M_n.$$

The value of a firm in the stationary equilibrium is $v_o = \frac{\pi_o}{\tilde{\delta}_{M_o}}$ with $o \in \{n, a\}$, where $\tilde{\delta}_{M_o} \equiv 1 - \nu(1 - \delta_{M_o})$ is the effective discount factor. Then, the free entry conditions in the stationary equilibrium are

$$\frac{WF_n}{(1 - G_n(\underline{Z}_n))} \frac{\tilde{\delta}_{M_n}}{\nu} = \bar{\pi}^n, \tag{4}$$

$$\frac{WF_a}{(1 - G_a(\underline{Z}_a))} \frac{\tilde{\delta}_{M_a}}{\nu} = \bar{\pi}^a. \tag{5}$$

The zero-profit condition in agriculture can be solved for the threshold, given optimal labor demand and the price of agriculture goods. Hence, equations 2, 3, 4 and 5 can be solved for the measure of farms in agriculture and firms in non-agriculture, as well as for the productivity of the marginal firms in each sector. This productivity determines the equilibrium distribution of firms operating in each sector. Wages and employment allocations are functions of these equilibrium quantities.

4 Quantitative Analysis

In this section, we present the quantification of our model.[17] 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 of 2p.p., 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 2.4. Finally, we evaluate the impact of the same shock under

[17] The current version presents steady state to steady state comparisons. Next version will characterize the full transition path.

Table 3: 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.

counterfactual scenarios where farms lack the ability to become more capital intensive as relative prices change.

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.[18]

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 $\gamma_l = 0.18$ 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

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

Table 4: 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.

so that the equilibrium labor share correspond to the calibrate one, given capital-output ratios.[19]

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 3 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 4.

[19] 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.

4.2 Comparative Statics or 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.

Figure 10: 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.

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 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 10 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

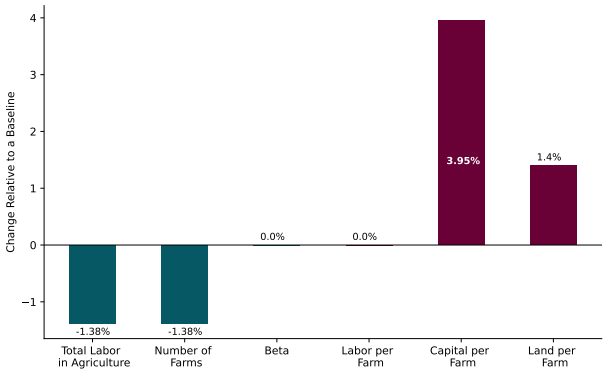
To understand the role of mechanization in shaping agricultural responses to non-agricultural shocks, we contrast two theoretical environments from our model: one in which mechanization is endogenous described in previous Section, and another where mechanization channel is shut down and the level of 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 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 11.

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

Figure 11: 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.

farm exits.

4.3 Counterfactuals *Work-in-progress*

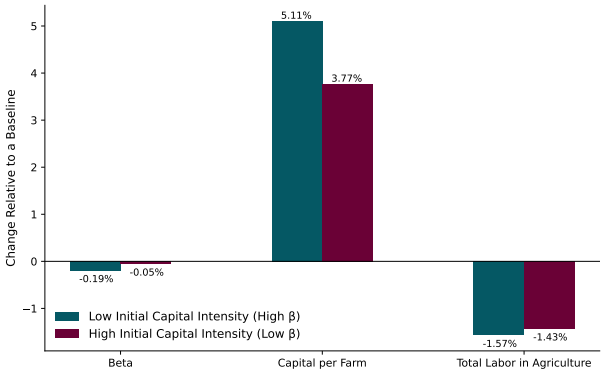
Building on our baseline counterfactual, we plan to extend our quantitative analysis by incorporating initial heterogeneity in agricultural organization across regions. Brazil exhibits substantial variation in the structure of agricultural production—particularly in the functioning of land and capital 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).[20] The main model outcomes for this higher-mechanization case are shown in Figure 12.

Figure 12: 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.

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.

[20] 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.

5 Conclusion

This paper explores the role of shocks to the non-agriculture sector for the organization of production in the agriculture sector. Using data from Brazil, we show that large-entry shocks into the non-agriculture sector are associated with labor reallocation out of agriculture.

These patterns are at odds with a common parameterization of the canonical model of structural change, which would predict labor reallocation into agriculture in response to entry and productivity improvements in non-agriculture. By allowing for endogenous capital intensification in the agricultural sector, we can reconcile the theory with the data.

For a calibrated economy to the Brazil experience, we document that labor demand shock consistent with an increases of real wages of 1.4% can account for a reallocation of labor out of agriculture of 1.38% and an increase in capital intensity of the farms of 5.11%. This mechanization channel is accompanied by land consolidation, through farm exit and 1.4% higher land per farm. In addition, large entry shocks in non-agriculture display heterogeneous effects across economies, depending on the level of capital intensity at baseline, which we will bring to the data observed across different municipalities in Brazil and quantitatively explore in the next version of this manuscript.

References

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

ACEMOGLU, D. AND F. ZILIBOTTI (2001): "Productivity Differences*," *The Quarterly Journal of Economics*, 116, 563–606.

ALVAREZ-CUADRADO, F. AND M. POSCHKE (2011): "Structural Change Out of Agriculture: Labor Push versus Labor Pull," *American Economic Journal: Macroeconomics*, 3, 127–58.

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., G. GARBER, AND J. PONTICELLI (2020): "Capital Accumulation and Structural Transformation*," *The Quarterly Journal of Economics*, 135, 1037–1094.

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.

CHEN, C. (2020): "Technology adoption, capital deepening, and international productivity differences," *Journal of Development Economics*, 143, 102388.

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.

FOSTER, A. D. AND M. R. ROSENZWEIG (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.

IMBERT, C. AND G. ULYSSEA (2023): "Rural Migrants and Urban Informality: Evidence from Brazil," Working Paper 18160, CEPR.

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

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

SURI, T. AND C. UDRY (2022): "Agricultural Technology in Africa," *Journal of Economic Perspectives*, 36, 33–56.

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.

Appendix

A Additional Tables and Figures

Figure A1: Epicenter versus Surroundings: Baseline urban share

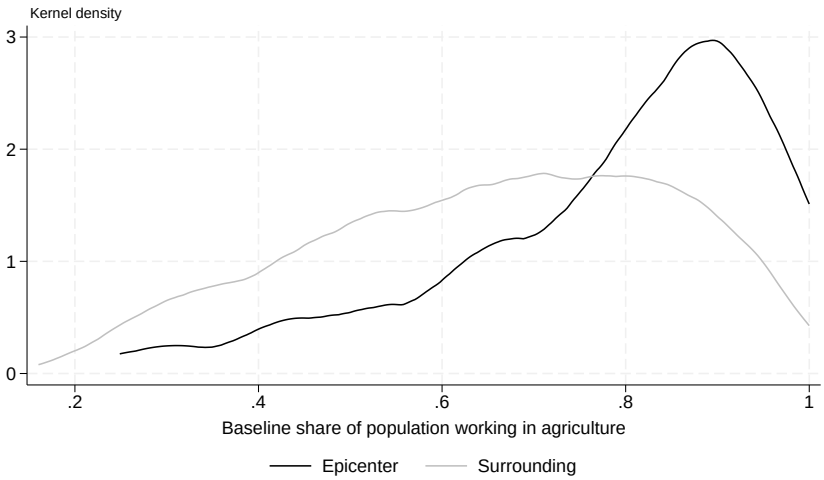
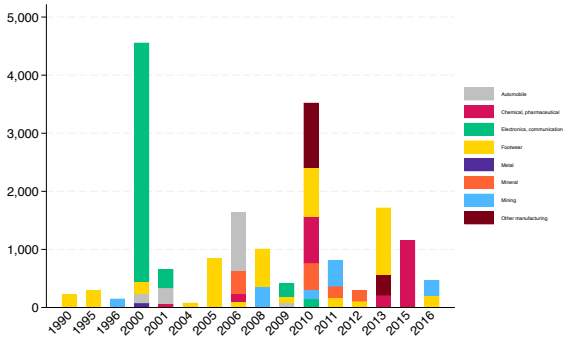
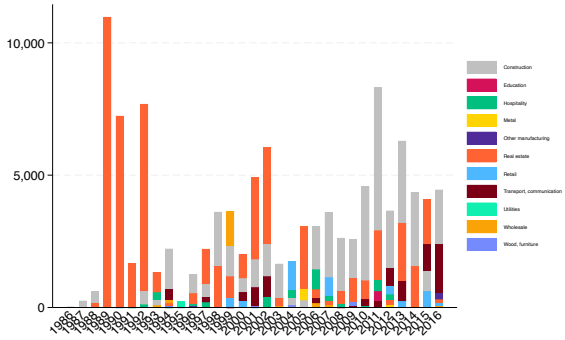


Figure A2: Total Employment from Large Entries at Entry Year

Panel A: Tradables



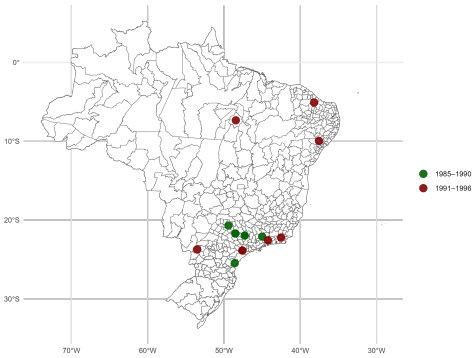
Panel B: Non-tradables



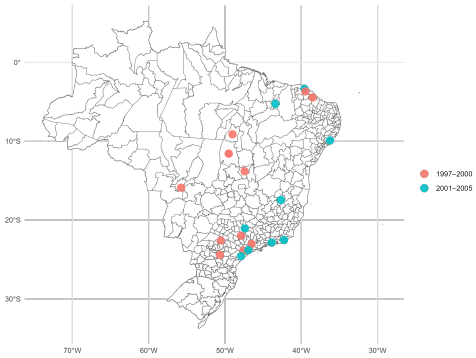
Notes: A large entry is a new non-agricultural establishment with more than 0.1% of total micro-region population at entry. Excludes entries in manufacturing activities that directly demand agricultural inputs (i.e., food, beverages, textiles, and paper).

Figure A3: Spatial Distribution of Large Entry Shocks

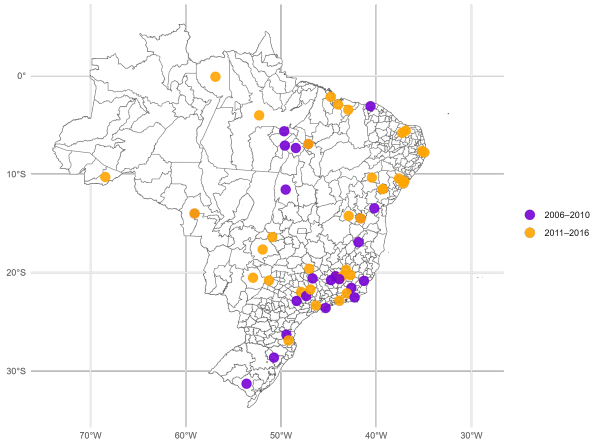
Panel A: 1985-1996



Panel B: 1997-2005



Panel C: 2006-2016



Notes: This figure shows the spatial distribution of large entry shocks. Borders show 486 micro-regions, which are sets of contiguous municipalities.

Figure A4: Epicenter versus Surroundings: Baseline urban share

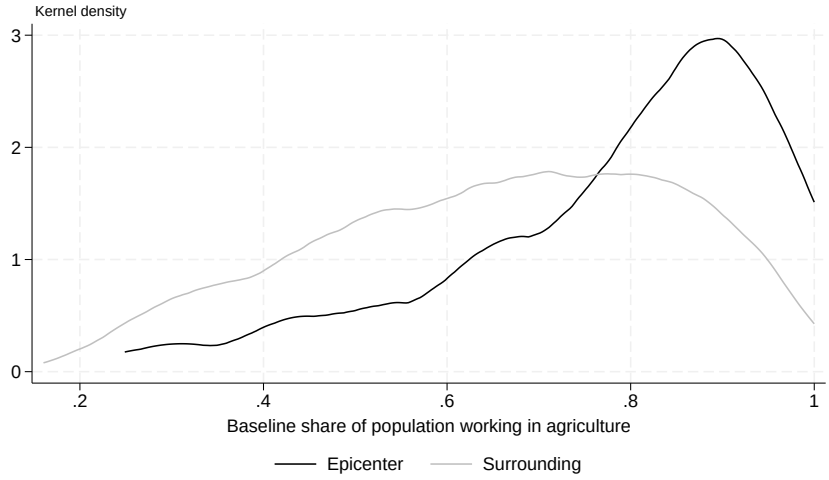
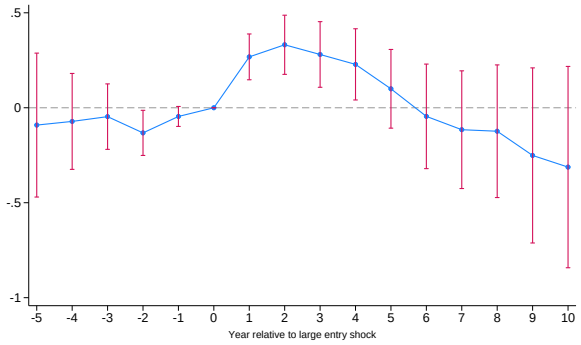
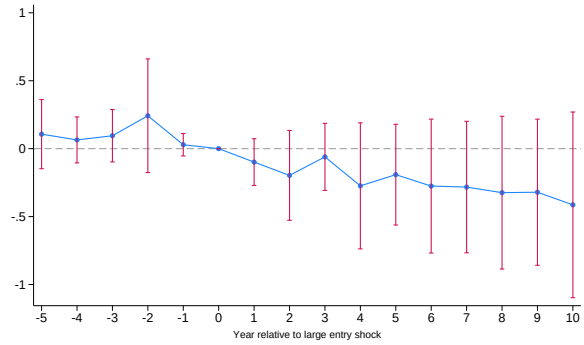


Figure A5: Large Firm Entry Shock on (at most) Primary Educated Workers by Sector

By Tradability

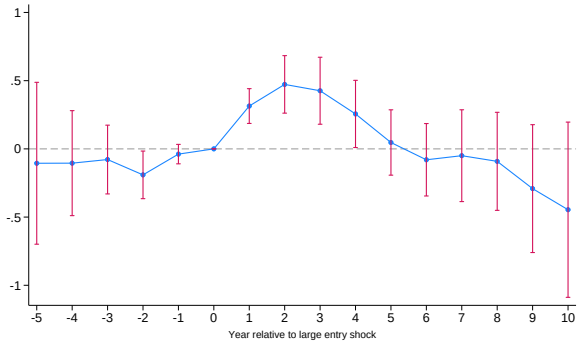


(a) Non-Tradable Industries

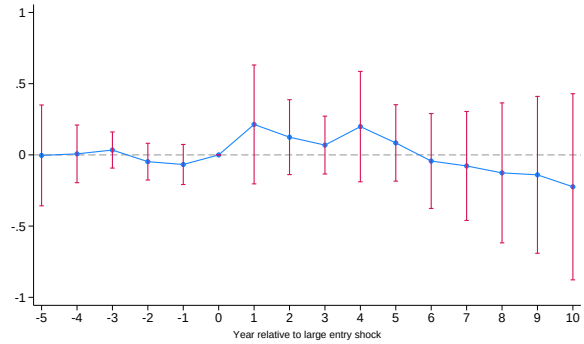


(b) Tradable Industries

(c) By Main Product



(d) Construction



(e) Real State

Notes:This figure presents difference-in-differences estimates of a large entry shock on epicenter municipality outcomes relative to surrounding municipalities. Estimates allow for heterogeneous treatment effects and use never-treated surrounding municipalities as controls, a la De Chaisemartin and d'Haultfoeuille (2020). A large entry is a new non-agricultural establishment with more than 0.1% of total micro-region population and at least 200 employees at entry. Excludes entries in manufacturing activities that directly demand agricultural inputs (i.e., food, beverages, textiles, and paper).

B Planning problem

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

Consider a planner that maximizes the Welfare

$$W_t \equiv \max_{\{C_{at}, C_{nt}, C_{at}\}} \sum_t \nu^t U(C_t)$$

for $\nu \in (0,1)$ and $U(C_t)$ Stone-Geary,

$$U(C_t) = (\theta_n \ln(C_{nt} - \bar{c_n}) + \theta_a \ln(C_{at} - \bar{c_a}) + \theta_s \ln(C_{st} - \bar{c_s}))$$

The planner takes technologies and the feasibility constrains of the economy as given.

$$Y_{n_i t} = Z_{n_i t} N_{nit}^{\alpha},$$

$$Y_{st} = Z_{st} N_{st}^{\alpha},$$

$$Y_{a_j t} = Z_{a_j t} \left(K_{jt}^{1-\beta} N_{a_j t}^{\beta}\right)^{\gamma} L_{jt}^{\gamma_L}.$$

$$M_n \int_{\underline{Z}_n}^{\infty} N_n(Z) g_n(Z) dZ + N_s + M_a \int_{\underline{Z}_a}^{\infty} N_a(Z) g_a(Z) dZ \equiv N^p,$$

$$N^f \equiv F_a m_a + F_n m_n + M_a \int_{\underline{Z}_a}^{\infty} f_a g_a(Z) dZ + M_n \int_{\underline{Z}_n}^{\infty} f_n g_n(Z) dZ,$$

$$N^p + N^f = \bar{N}. \qquad (\mu_N)$$

$$M_a \int_{\underline{Z}_a}^{\infty} K_{jt} g_a(Z_j) dZ_j \equiv K_t. \qquad (\mu_k)$$

$$\bar{N} l = M_a \int_{\underline{Z}_a}^{\infty} L_{aj} g_a(Z_j) dZ_j = M_a L_a. \qquad (\mu_L)$$

$$C_a = M_a \int_{\underline{Z}_a}^{\infty} Y_{aj} g_a(Z_j) dZ_j = M_a Y_a,$$

$$C_s = Y_s, \qquad C_n + X_t = M_n \int_{\underline{Z}_n}^{\infty} Y_{ni} g_n(Z) dZ = M_n Y_n. \qquad (\lambda_o, o \in \{a, s, n\})$$

$$K_{t+1} = X_t + (1 - \delta_k) K_t. \qquad (\mu_X)$$

$$M_{at+1} = m_{at}(1 - G(\underline{Z_{at+1}})) + M_{at}(1 - \delta_{M_a}), \qquad (\mu_a)$$

$$M_{nt+1} = m_{nt}(1 - G(\underline{Z_{nt+1}})) + M_{nt}(1 - \delta_{M_n}). \qquad (\mu_N)$$

The planner's problem is to maximize the present discounted value of welfare given technologies by choosing the measure of firms (and entrants) in each sector, the threshold productivity for active firms, the level of consumption, investment and capital available next period, as well as the allocation of employment across firms, the allocation of capital and land across farms. The optimality conditions associated to this problem are

$$\mu_{at}(1 - G(\underline{Z_{at+1}})) = F_a \mu_{Nt}, \tag{m_{at}}$$

$$\mu_{at} = \nu \left[\pi_{at+1} + \mu_{at+1}(1 - \delta_{M_a}) \right], \tag{M_a}$$

where

$$\begin{aligned} \pi_{at+1} \equiv \; & \lambda_{at+1} \int_{\underline{Z}_{at+1}}^{\infty} Y_{aj} g_a(Z_j) dZ_j - \mu_{nt+1} \int_{\underline{Z}_{at+1}}^{\infty} N_{aj} g_a(Z_j) dZ_j - f_a \mu_{nt+1} - \\ & \mu_{Kt+1} \int_{\underline{Z}_{at+1}}^{\infty} K_{jt} g_a(Z_j) dZ_j - \mu_{Lt+1} \int_{\underline{Z}_{at+1}}^{\infty} L_{aj} g_a(Z_j) dZ_j \end{aligned}$$

corresponds to average profitability from a new farm in the market.

$$\mu_{Nt}(1 - G(\underline{Z_{nt+1}})) = F_n \mu_{Nt}, \tag{m_{nt}}$$

$$\mu_{Nt} = \nu \left[\bar{\pi}_{nt+1} + \mu_{nt+1}(1 - \delta_{M_n}) \right], \tag{M_n}$$

where

$$\bar{\pi}_{nt+1} \equiv \lambda_{nt+1} \int_{\underline{Z}_{nt+1}}^{\infty} Y_{ni} g_n(Z_i) dZ_i - \mu_{nt+1} \int_{\underline{Z}_{nt+1}}^{\infty} N_{ni} g_n(Z_i) dZ_i - f_n \mu_{nt+1}$$

corresponds to average profitability from a new firm in non-agriculture.

The choice of investment and capital satisfies

$$\lambda_n = \mu_X, \tag{X}$$

$$\mu_{Xt} = \nu \left[\mu_{Xt+1}(1 - \delta_k) + \mu_{kt+1} \right], \tag{K_{t+1}}$$

with $\mu_{kt} = \lambda_{at} \frac{\partial Y_{a_j t}}{\partial K_{jt}}$.

On the margin, the productivity of the active firm satisfies

$$M_{nt+1} \nu \pi_{nt+1}(\underline{Z}_{nt+1}) = m_{nt} \mu_{Nt} g(\underline{Z}_{nt+1}), \tag{$\underline{Z}_{nt+1}$}$$

where $\pi_{nt}(\underline{Z}_{nt}) \equiv \lambda_{nt} Y_n(\underline{Z}_{nt}) - \mu_N N_n(\underline{Z}_{nt}) - f_n \mu_{Nt}$.

$$M_{at+1} \nu \pi_{at+1}(\underline{Z}_{at+1}) = m_{at} \mu_{at} g(\underline{Z}_{at+1}), \tag{$\underline{Z}_{a_t+1}$}$$

where $\pi_{at}(\underline{Z}_{at}) = \lambda_{at} Y_a(\underline{Z}_{at}) - \mu_N N_a(\underline{Z}_{at}) - \mu_K K_a(\underline{Z}_{at}) - \mu_L L_a(\underline{Z}_{at}) - f_a \mu_{Nt}$.

Notice that $g(\underline{Z}_{it+1}) = 0$ by definition, and therefore the zero profit condition boils down to $\pi_{it}(\underline{Z}_{it}) = 0$. Using the optimal labor demand in each sector one can write the profits as a constant fraction of revenue as follows

$$(1-\alpha)\lambda_{nt} Y_n(\underline{Z}_{nt}) = f_n \mu_{Nt},$$

$$(1-\gamma-\gamma_L)\lambda_{at} Y_a(\underline{Z}_{at}) = f_a \mu_{Nt}.$$

We can rewrite these conditions as a function of the average firm in each sector as

$$(1-\alpha)\lambda_n Z_n N_n^{\alpha} \left(\frac{\underline{Z}_n}{Z_n}\right)^{\frac{1}{1-\alpha}} = f_n \mu_N, \qquad \text{(zero profit n)}$$

$$(1-\gamma-\gamma_L)\lambda_a K_a^{(1-\beta)\gamma} N_a^{\beta\gamma} L_a^{\gamma_L} Z_a \left(\frac{\bar{Z}_a}{Z_a}\right)^{\frac{1}{1-\gamma-\gamma_L}} = f_a \mu_N, \qquad \text{(zero profit a)}$$

where $\bar{\pi}_n = (1-\alpha)\lambda_n Z_n N_n^{\alpha}$. Combining the optimality conditions for entry and the measure of active firms in each location yields the free entry condition in steady state, i.e.

$$\bar{\pi}_n = \frac{\tilde{\delta}_{M_n}}{\nu} \frac{F_n \mu_N}{(1 - G(\underline{Z}_n))}, \qquad \text{(free entry)}$$

where $\tilde{\delta}_{M_n} = (1 - \nu(1 - \delta_{M_n}))$ and as $\nu \to 1$, then $\tilde{\delta}_{M_n} \to \delta_{M_n}$.

The static allocation of employment across sectors, employment in non-agriculture, and employment capital and land in the agriculture sector are standard.

Since capital labor ratios, and land to labor ratios are equalized across farms, the relative employment allocation of farms is

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

The relative employment allocation in non-agriculture satisfies

$$\frac{Z_{ni}}{Z_{ni'}} = \left(\frac{N_{ni}}{N_{ni'}}\right)^{1-\alpha}.$$

To characterize the relative employment allocation across non-agricultural sectors, it is useful to define the average productivity in non-agriculture as $Z_n \equiv \left(\int_0^{\infty} Z_{n_i}^{\frac{1}{1-\alpha}} g(Z_{n_i}) dZ_{n_i}\right)^{1-\alpha}$. This average productivity is independent of the measure of firms in the market and depends on the conditional probability of each productivity level, g, which is in turn a function of the threshold productivity level in the sector.

Relative employment is

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

where N_n is the employment of the average firm in non-agriculture, i.e. $N_n = \left(\int_0^\infty N_{n_i}(Z_{n_i}) g(Z_{n_i}) dZ_{n_i}\right)$.

We can characterize the relative employment allocation between agriculture and non-agriculture in the surrounding areas as a function of labor productivities,

$$\frac{Y_a/N_a}{Y_s/N_s} = \frac{\alpha}{\gamma\beta}\frac{\lambda_s}{\lambda_a}.$$

where Y_a/N_a is the labor productivity of the average farm in the agriculture sector.

Finally, the optimality condition with respect to consumption in non-agriculture and in the surroundings are

$$\lambda_s = \frac{\theta_s}{Z_s N_s^\alpha - \bar{c}_s} \text{ and } \lambda_n = \frac{\theta_n}{M_n(1-s) Z_n N_n^\alpha - \bar{c}_n}. \tag{6}$$

Replacing these prices into the optimality condition for labor yields

$$\frac{Z_s}{Z_n} = \left(\frac{N_s}{N_n}\right)^{1-\alpha} \frac{\theta_n}{\theta_s} \frac{Z_s N_s^\alpha - \bar{c}_s}{M_n(1-s) Z_n N_n^\alpha - \bar{c}_n} \tag{7}$$

And similarly for agricultural labor using FOCs for labor,

$$N_a^{1-\gamma\beta} = \frac{\gamma\beta}{\alpha}\frac{\lambda_a}{\lambda_n}\frac{Z_a K_a^{(1-\beta)\gamma} L_a^{\gamma_L}}{Z_n N_n^{\alpha-1}}.$$

C Computational Algorithm

C.1 Algorithm

1. Use equations free entry and zero profit n (to eliminate average profits) to compute the threshold productivity in each sector. For example, in non-agriculture

$$f_n \left(\frac{Z_n}{\underline{Z}_n} \right)^{\frac{1}{1-\alpha}} = \frac{\tilde{\delta}_{M_n}}{\nu} \frac{F_n}{(1 - G(\underline{Z}_n))}$$

   Guess the price in non-agriculture, λ_n, and in agriculture, λ_a.

2. Use the Euler equation for capital in steady state to compute capital output ratio. then use this value in the price of agriculture from the optimality condition for agricultural consumption to solve for $M_a K_a$.

$$1 = \nu \left[(1 - \delta_k) + \frac{\lambda_a}{\lambda_n} (1 - \beta) \gamma \frac{Y_a}{K_a} \right]$$

$$\lambda_a = \frac{\theta_a}{M_a K_a \frac{Y_a}{K_a} - \bar{c}_a}$$

3. Use the law of motion of capital in the steady state to get an expression for the level of investment as a function of the stock of capital $X = \delta_K M_a K_a$.

4. Use the optimality condition for consumption in the non-agriculture sector to infer the savings rate

$$s = \left(\frac{\frac{\theta_n}{\lambda_n} + \bar{c}_n}{\delta_k M_a K_a} + 1 \right)^{-1}$$

5. Use the free entry and the optimality condition for labor in non-agriculture

$$\mu_N = \alpha \lambda_n Z_n N_n^{\alpha - 1}$$

$$(1 - \alpha) \lambda_n Z_n N_n^{\alpha} = \frac{\hat{\delta}}{\nu} \frac{F_n \mu_N}{1 - G(\underline{Z}_n)}$$

   so

$$N_n = \frac{\hat{\delta}}{\nu} \frac{F_n}{1 - G(\underline{Z}_n)} \frac{\alpha}{1 - \alpha}$$

6. Infer the wage level from the optimality condition in non-agriculture,

$$\mu_N = \alpha \lambda_n Z_n N_n^{\alpha-1}$$

and then use the optimality condition for consumption in the agriculture sector plus the free entry to infer the measure of firms in agriculture.

$$M_a = \frac{(\theta_a + \bar{c}_a \lambda_a)(1-\gamma-\gamma_l)}{\frac{\delta}{\nu}\frac{F_a \mu_N}{1-G(\underline{Z}_a)}}$$

7. Use this measure of firms, plus the value of $M_a K_a$ to infer the level of capital in steady state.

8. Use the optimality condition in non-agriculture

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

9. Use the labor optimality in the agriculture and the epicenter, plus the level of capital to infer the labor in agriculture,

$$N_a^{1-\gamma\beta} = \frac{\gamma\beta}{\alpha}\frac{\lambda_a}{\lambda_n}\frac{N_n}{Y_n} Z_a {K_a}^{\gamma(1-\beta)} L_a^{\gamma_l}$$

10. Use the ratio of the labor optimality conditions in non-agriculture and the surroundings to infer

$$\frac{Z_s}{Z_n} = \left(\frac{N_s}{N_n}\right)^{1-\alpha} \frac{\lambda_n}{\theta_s}(Z_s N_s^{\alpha} - \bar{c}_s)$$

11. Use the optimality condition for consumption in the surroundings to infer the price

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

12. Go back to the optimality condition of consumption in the agriculture sector to infer its price

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

this requires computing agriculture output which we can do once we have employment in step 9.

13. Use the labor market clearing condition to update the guess of the price in non-agriculture

$$\overbrace{M_n N_n + N_s + M_a N_a + F_a \frac{\hat{\delta}}{\nu} \frac{M_a}{1 - G(\underline{Z}_a)} + F_n \frac{\hat{\delta}}{\nu} \frac{M_n}{1 - G(\underline{Z}_n)} + f_a M_a + f_n M_n}^{\text{LHS}} = \bar{N}$$

If LHS$> \bar{N}$ we need higher wages and higher price of output in non-agriculture λ_n. If RHS$< \bar{N}$ we should lower the price in non-agriculture.