Commodity Booms and Structural Transformation: The Role of Input Use and Land Inequality

Andrei Arminio Laskievic*

May 2021

Abstract

We study the effects of an agricultural commodity price boom on structural transformation. We construct a shift-share measure of exposure to the commodity shock by combining climate- and soil-induced variation in agricultural production patterns among municipalities in Brazil with fluctuations in international commodity prices between 2000 and 2010. We show that labor was reallocated away from agriculture and into the manufacturing sector in locations more exposed to the commodities boom. Using data from the Population and Agricultural Censuses, we argue that the results are consistent with greater use of capital inputs in agriculture, that worked as substitutes for farm labor, and greater land inequality in more exposed locations, which ultimately displaced workers from the agricultural sector.

*São Paulo School of Economics–FGV. E-mail: alaskievic@gmail.com

1 Introduction

The relationship between agricultural endowments and economic development has been a long-studied subject in the growth literature (Engerman and Sokoloff, 1994; Vollrath, 2011). Whether the expansion of the agricultural sector promotes or curbs structural transformation is a fundamental and highly debated question on development economics (Lewis, 1954; Rostow, 1960; Gollin et al., 2002). In particular, open economies can experience both an increase or decrease in industrialization depending on the comparative advantage of the agricultural sector and its export composition (Matsuyama, 1992). Moreover, the classical theoretical models have focused on productivity changes, both in agriculture and manufacturing, as the main driver of structural change. The bulk of the studies have modeled the transition from the Malthusian era to the modern economic growth regime, but the effects of the agricultural sector expansion on structural change might depend on an economy's stage of development (Hansen and Prescott, 2002; Ngai and Pissarides, 2007).

Few studies have explored the structural transformation mechanisms in modern or developing economies. While shocks to agricultural productivity have been a recent subject of study in empirical works, the evidence regarding its effects and possible mechanisms on modern local economies is still scarce and ambiguous (Bustos et al., 2016; Moscona, 2018; Uribe-Castro, 2019). Such studies also tend to focus on particular crops and might miss the general effects that impact the agricultural sector more generally. Given the existing ambiguity in both theoretical and empirical studies, the direction of the change in structural transformation is context-specific and depends on the type of shock to the economic structure. Isolating the forces that can operate such relationship is a fundamental challenge in the literature and is, ultimately, an empirical question.

We explore the 2000s commodities boom as an exogenous economic shock and analyze how the underlying structure of the agricultural sector in Brazil changed after the sharp rise in the prices of a large basket of internationally traded commodities. In particular, we look at the change in employment shares across sectors and argue that differential input use (labor, capital, and land) worked as a key mechanism in driving the effects. Additionally, we explore inequality in landholdings as another important mechanism in generating structural change. Land inequality is often overlooked in the structural transformation literature and can be important in accounting for both changes in agricultural productivity and the displacement of workers.

To empirically assess the effect of the commodity boom on local development and structural change, we construct a shift-share measure to isolate municipalities that were more exposed to the shock. We construct the exposure shares by using actual observed yields in

Brazilian municipalities and data on crop potential yields from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) project. We combine both variables into a fractional multinomial logit (FML) model of crop choice, which calculates predicted shares for the crops in our sample for each municipality. We then interact the shares with the exogenous variation in international prices of commodities between 2000 and 2010, which represents our shift (or shock) variable.

We first explore the effect of the commodity boom on employment shares and find that regions that were more exposed to the commodity boom experienced a reallocation of labor away from agriculture and into the manufacturing sector. Our estimates imply that the commodity boom can explain 10% of the observed differences in the reduction of the agricultural share across Brazilian municipalities and 18% of the corresponding differences in the growth of the manufacturing employment share. We also do not observe any effect of the commodities shock on employment share in services at the local level.

Next, we explore potential mechanisms that led to the observed reallocation of workers. First, we explore how the intensity in inputs use changed in more affected regions. We observe an increase in the use of capital, measured by the use of machines and genetically engineered seeds in the farms. We argue that the higher use of capital in the farms worked as substitutes to agricultural workers. We then explore how the change in land inequality after the shock might have contributed to the displacement of agricultural workers. We find that land inequality increased in more exposed municipalities by looking at the Land Gini and the number and area of the largest farms inside the regions. The share of land used in agriculture production was disproportionately appropriated by a small number of large properties, which displaced small rural owners and workers from the agricultural sector.

Our main findings are robust to a series of deviations from our baseline framework. We first explore if our main results hold when looking only at formal jobs from a matched employer-employee dataset and find that the general conclusions remain unaltered. Next, we use a series of different definitions for our exposure measure and also find similar results. Our main estimates also remain statistically significant when we correct standard errors to account for spatial correlation and correlation across the exposure shares. Finally, our results are robust to using microregions as a larger unit of observation. We then conclude that the 2000s commodities boom resulted in structural transformation. The use of capital inputs and rising land inequality increased in agriculture and labor was reallocated away from the agriculture sector and into manufacturing, spurring industrial growth.

2 Related Literature

Resource Booms and Development: Commodity price shocks exogenous to the local economies are often pointed out as an important force behind a resource curse (Prebisch, 1950; Singer, 1950). The hypothesis, usually called the Prebisch-Singer hypothesis, states that agricultural commodity exports would hurt the manufacturing sector and curb long term economic growth via declining terms of trade and dependency on the primary sector.[1] Such discussion is especially prominent in countries that rely heavily on the agricultural sector and that have experienced high rates of growth during the 2000s commodities supercycle, which also led to an observable expansion of activities associated with the primary sector, often called the deindustrialization of the economy—or rather, "reprimarization".

The first studies that tried to evaluate the resource curse empirically—using mainly cross-sectional country data—often found that countries more affected by resource booms did not experience long-term growth gains (Sachs and Warner, 1999, 2001). A large body of the literature has also focused on political factors that might worsen or alleviate the resource curse. In particular, diverging patterns of development after the booms might be explained by the quality of local institutions (Mehlum et al., 2006).[2] But another recent strand of the literature has focused more on robust microeconometric evidence using within-country data and found that resource booms can actually generate manufacturing growth via different mechanisms. The empirical strategy and economic context of our study are related to some of such studies.

Allcott and Keniston (2018) revisit the question regarding the existence of a resource curse in the context of the discoveries of natural resources and their effect on local economies in the United States. By following a shift-share approach and interacting each county's cross-sectional variation in initial oil and gas endowment with the time series variation in national employment, the authors study the effects of oil booms and busts on local wages and employment—while also disentangling the agglomeration and crowding-out effects of the booms—and find that both wages and manufacturing employment increase during the booms. Also, they do not find any long-run effects on productivity decline, implying that a Dutch Disease mechanism did not occur in their setting.

Exploring the Brazilian context, Cavalcanti et al. (2019) show how oil discoveries increased local production and had spillover effects over other sectors of the economy. Workers were reallocated away from agriculture and into manufacturing and services, increasing urbanization. Dube and Vargas (2013) explore how income shocks, arising from increas-

[1] For a recent discussion, see Harvey et al. (2010).

[2] See Van der Ploeg (2011) and Van Der Ploeg and Poelhekke (2017) for a comprehensive discussion about the Dutch Disease literature.

ing commodities prices, affect civil conflict in Colombia. Their empirical strategy interacts commodity production for coffee and oil across different municipalities with their respective international prices. The results show important heterogeneous impacts of the type of commodity on the incidence of political violence: intensification of attacks are associated with a declining price of coffee and also with a rise in the prices of oil. The channels for the different impact come from the relative effects of both coffee and oil on local wages and revenues. In a historical setting, Uribe-Castro (2019) studies the same Colombian agricultural sector during the 20th century and finds that coffee price shocks impacted cultivation and increased the opportunity cost of education, which reduced the supply of skilled workers, and slowed down structural transformation.

Agricultural Productivity and Structural Change: In a series of works starting with Bustos et al. (2016), the authors show that the adoption of new agriculture technologies can affect structural change in different directions. In particular, they explore the introduction of a new variety of genetically engineered (GE) soybean seed and of a second season in maize cultivation in Brazil. They argue that the GE soy seeds represent a labor-saving technical change in agriculture, while the new maize variety can be considered a labor-augmenting technology. The difference in labor-intensity is fundamental for the effects on structural change. In the case of a labor-saving technology, labor is reallocated away from agriculture and into manufacturing, while the opposite effect happens with the labor-augmenting technology. In a subsequent work, Bustos et al. (2020a) show that the adoption of GE soy generated an increase in savings that were reinvested in urban regions and not in rural areas. Capital, therefore, has also reallocated towards the manufacturing sector and urban areas.

In another study, Bustos et al. (2020b) show that the previous reallocation of labor away from agriculture reinforced comparative advantage in the least skill-intensive manufacturing industries. Such effect happened because the displaced workers were relatively unskilled and had lower levels of human capital. The findings lead to an important conclusion about long-run economic growth: it is often assumed that structural change towards manufacturing will increase the overall total productivity of the economy, but if it is an unskilled-bias reallocation, the least R&D-intensive manufacturing industries will absorb most of the labor, decreasing the size of the skill-intensive industries and slowing down innovation and manufacturing productivity growth.

While the series of works have deepened our knowledge about agricultural productivity and structural change, it still depends on a particular type of commodity. There is also evidence that other types of technological changes in other agricultural crops imply different effects on trends of structural change. Moscona (2018) shows that productivity growth associated with the Green Revolution in the last decades has been complementary to labor, increasing the employment share in agriculture both within and across countries, while also reducing urbanization and manufacturing employment.

2000s Commodities Supercycle and the Brazilian Labor Market: Some recent works have also explored the Brazilian economy as a setting for studying the effects of the commodity boom on labor market outcomes. Costa et al. (2016) use the "China shock" approach to show which sectors were benefited or harmed by the increase in Chinese imports supply and exports demand. The authors find that the wages of agricultural and extractive sector were positively affected, while the manufacturing sector was negatively affected. Benguria et al. (2018) show that higher commodity prices in Brazil increase the domestic demand and induce wage increases in unskilled-intensive industries. Adão (2015) finds that movements in world commodity prices explain part of the decrease in Brazilian wage inequality from 1991–2010.

Shift-Share Instruments: Related to our empirical identification strategy, a new body of the literature has further formalized the shift-share econometric approach pioneered in Bartik (1991) and Blanchard et al. (1992). Some of them provide very useful ways of implementing the shift-share design in our study. A priori, one must decide from which part of the instrument the exogenous variation will come from, the shocks or the shares. Goldsmith-Pinkham et al. (2020) provide a framework for working with exogeneity coming from the shares. On the other hand, Borusyak et al. (2018) and Borusyak and Hull (2020) offer a setting for working with quasi-random assignments of shock, while allowing the exposure shares to be endogenous. The framework is more in line with our strategy since we provide a set of international price shocks that are plausibly uncorrelated with other unobserved factors affecting local outcomes.

Theoretical Framework: We ground our analysis on the theoretical literature that explores the relationship between agricultural development and structural change in open economies. The literature is relatively scarce in comparison to the bulk of studies that analyze structural change in closed economies. We will then focus on the particular models that better approximates our scenario. Gollin et al. (2016) develop a framework in which urbanization can occur in economies even without industrialization. In their model, an exogenous increase in resource export earnings raises incomes, which then causes an increase in the urbanization rate as the rural sector contracts. Their results imply that a resource boom will ultimately generate an urbanization pattern just as effectively as industrial development, creating "consumption cities" as opposed to "production cities".

Corden and Neary (1982) construct a three-sector open economy model with traded and nontraded goods. After a boom in one of the traded sectors, which is considered to be extractive, they explore the structural change effects based on the resource movement effect and the spending effect. The resource movement effect raises the marginal products of the mobile factors after the boom and draws resources out of the other two sectors. The spending effect raises the spending on the service sectors after the rise in income generated by the boom. The overall effect on industrialization depends on the relative strength of the two effects and on the mobility of factors across sectors.

3 Motivating Facts

The Brazilian economy relies heavily on agriculture. For example, Brazil is among the largest producers in the world of coffee, sugarcane, soybean, maize and beef. Furthermore, it is also characterized by regional variations in the degree of specialization in commodity output. The production of the agricultural crops, together with cattle raising, provide the basis for the large agribusiness industry in Brazil, which represents 22% of Brazil's GDP, a third of its employment, and almost 40% of its exports (PwC, 2013; FGV, 2015).

Products derived from agricultural production, like orange juice, ethanol, and paper also represent an important dimension of the large agribusiness industry in Brazil. In particular, the production of the derived products is characterized by high use of technology and agroindustry techniques. The agroindustry is, therefore, another important dimension in our analysis of the agricultural expansion, since part of the agricultural sector in Brazil can be characterized by economies of scale that are normally associated with the manufacturing sector. Furthermore, the agroindustry captures both farming production and the supply of farming inputs such as machinery and seeds.

Brazil has also one of the highest levels of land inequality in the world (Bauluz et al., 2020). According to the last Agricultural Census in 2017, less than 1% of all farms concentrate about half of the total rural area. The concentration of land remains surprisingly constant over time. Ever since the first Agricultural Census in 1920, the number of rural properties with more than 1000 hectares in area varied between 4% and 0.7% of the total number of agricultural establishments, but appropriated between 35% and 63% of the total area devoted to agriculture over the last 100 years. Figure 1 illustrates the differences in the Land Gini for the Brazilian municipalities in 2017.

The characteristics of the Brazilian agricultural sector presented so far make the country a suitable setting to analyze the dynamics of structural change after the 2000s commodities boom. The last commodity supercycle started in 2000 and peaked in around 2011. It's magnitude is comparable to the other two historically observed supercycles that followed the Second Industrial Revolution and the aftermath of the Second World War (Erten and Ocampo, 2013). Figure 2 illustrates how the prices of non-mineral agricultural commodities

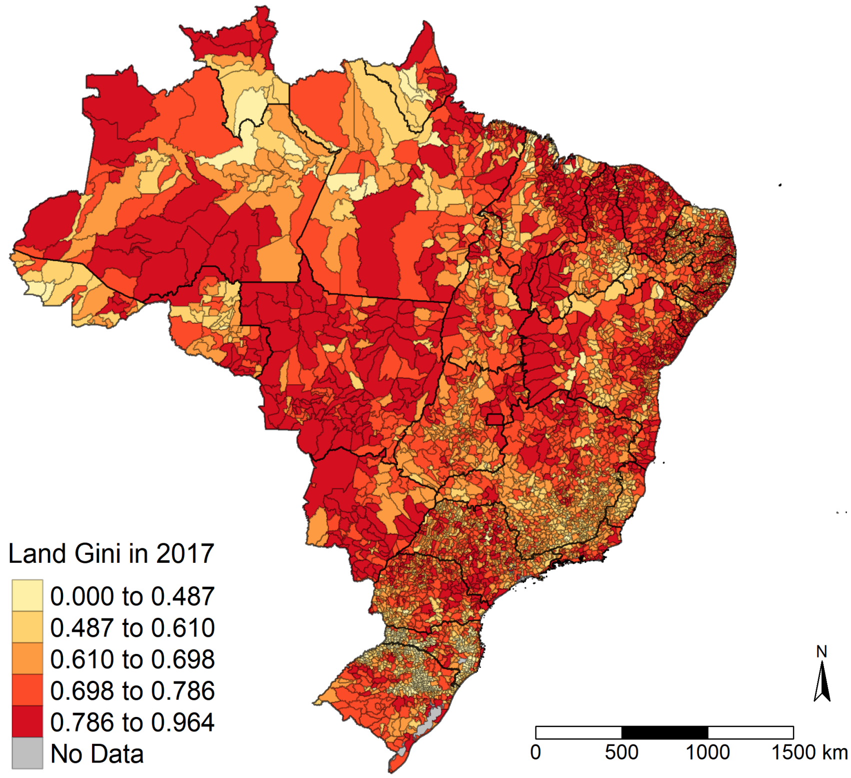


Figure 1: Land Inequality in Brazil

Notes: Author's calculations using the Brazilian Agricultural Census data from 2017. See Appendix A.1 for variable definition and sources and Appendix A.2 for details about the calculation of the Land Gini.

peaked very rapidly in about ten years after the 1990 decade when they were comparatively low.

Brazilian municipalities have been experiencing a long structural transformation since the 1960s, when the economy started to grow and industrialize rapidly. Figure 3 shows the evolution of employment shares in the main sectors of the Brazilian economy over time. Although the structural change has consolidated in the 1980 decade, we still observe a substantial reallocation of workers away from agriculture in the last decades, especially between 2000 and 2010. In particular, the distribution of the employment shares differs substantially when analyzed at the municipality level.

The facts discussed above call for a careful examination of the differential trends of local development in the municipalities. We, therefore, analyze how the commodity boom affected the structure of the agriculture sector and regional development in the Brazilian economy by paying special attention to the highlighted characteristics and facts.

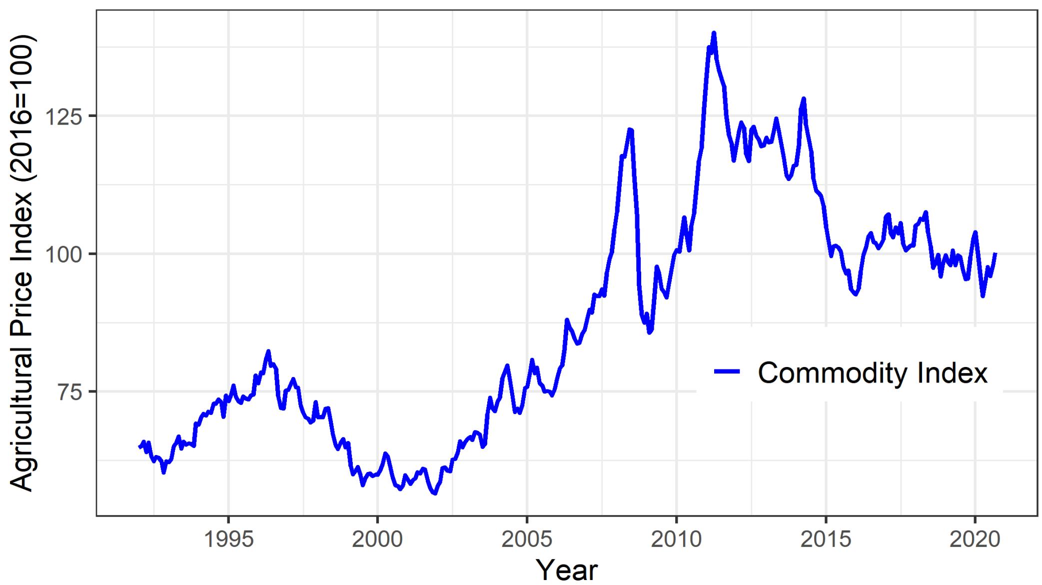


Figure 2: 2000s Commodities Supercycle

Notes: Commodity Price Index in US$ for agricultural products. Data from the World Bank's Global Economic Monitor. See Appendix A.1 for variable definition and sources.

4 Data

To construct our shift-share measure of exposure to the 2000s commodities supercycle, we first obtain information on the local municipalities production of agricultural crops from the Instituto Brasileiro de Geografia e Estatística (IBGE). More specifically, we use the Produção Agrícola Municipal (PAM), which is held annually at the municipality level and has information on a set of variables such as planted area, quantities and production values for a large set of crops produced in Brazil.

To obtain beef quantities, we also use the Pesquisa da Pecuária Municipal (PPM), which is similar to the PAM and accounts for the creation of livestock in Brazil. The only drawback is that the PPM reports the number of cattle heads, not the actual weight of beef produced. To overcome such problem, we multiply the total number of livestock in each municipality by 0.23, which is the average total carcass weight at slaughter in metric tons, to obtain a proxy for the production of beef in each municipality.[3]

We use data from the FAO-GAEZ project to obtain potential yields for each crop in our sample. The database consists of maximum attainable yields for multiple crops (in tons per hectare per year), which are considered as soil- and climate-based potential yields or

[3]The average total carcass weight at slaughter in Brazil comes from the Pesquisa Trimestral do Abate de Animais from IBGE.

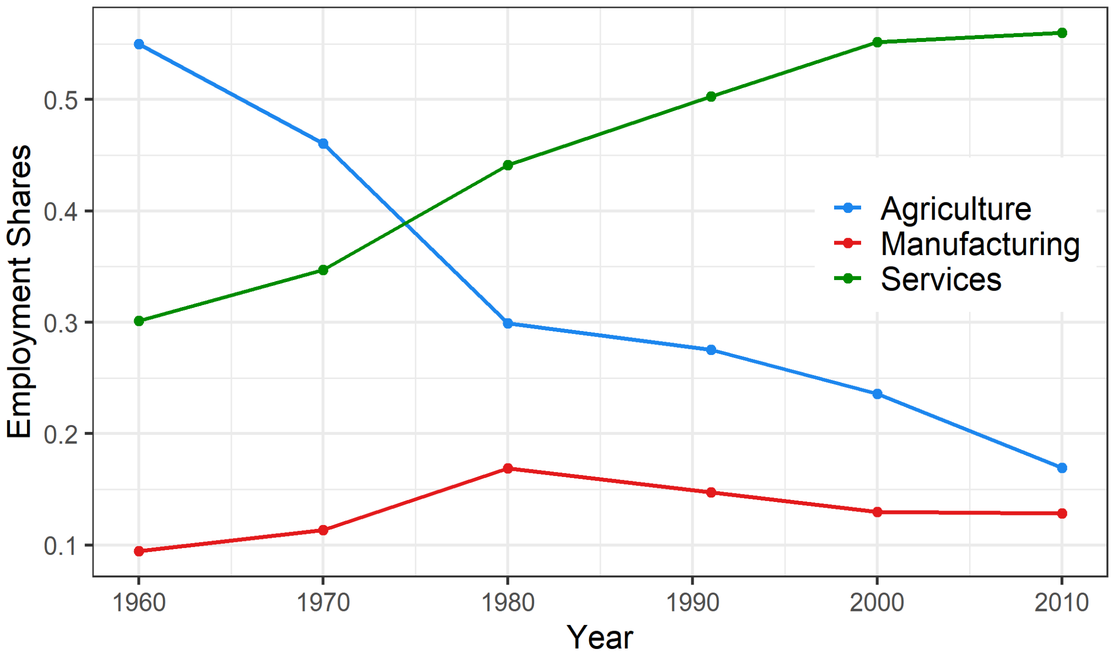


Figure 3: Employment Shares Across Sectors in Brazil

Notes: Author's calculations from IPUMS International microdata for Brazilian Population Censuses. See Appendix A.1 for variable definition and sources.

productivity (total production capacity) measures. The crop-specific data relies purely on exogenous geo-climatic features, such as local soil and weather characteristics, which are then incorporated into a model that predicts the maximum attainable yields for each crop in a given area under different choices of water supply methods and input levels of technology use.

The dataset is reported on a grid cell level, with each grid representing a 5-arc-minute resolution. At the 5-arc-minute level, there are around 255,680 grid cells in Brazil, representing around 11% of all grid cells on earth excluding bodies of water and ice shelves (Costinot et al., 2016). The area of the cells ranges from 8,586 hectares at the equatorial level to 6,031 hectares at the southern-most location in Brazil.

We aggregate the potential yields for each crop at the local level taking the mean value of each grid cell that falls within the border of each municipality and match all the crops present in the FAO-GAEZ with the previously obtained from the PAM/PPM dataset. We use data for rain-fed conditions under intermediate input technology—defined as a system in which "production is based on improved varieties, on manual labor with hand tools and/or animal traction and some mechanization".[4] As a robustness check, we propose to analyze the

[4]For more details about the documentation of the FAO-GAEZ project, see Appendix A.1 and IIASA/FAO

sensitivity of our results to using the high inputs.[5] With both PAM/PPM and FAO-GAEZ, we can then construct our "share" measure, consisting of the total production of crops in each municipality averaged over the 5 years preceding the start of our analysis.[6]

For the "shock" part of our measure, we use the international commodity prices from the Global Economic Monitor (GEM) Commodities database prepared by the World Bank, which contains monthly information over an extended period on the price of several commodities that are traded in the international markets. We then average the prices over the twelve months within each year to obtain yearly values and standardized each price to US dollars per ton.[7] In some cases, there may be a single price that matches multiple crops in the quantities dataset, so we consolidate the shift-share measure by matching each possible crop to its broad international price, leaving space to more than one commodity being assign to a single price[8], while dropping the cases in which we cannot establish a match at all.

For the main outcomes and controls, we use data from the Brazilian population census, the agricultural census, and Ipeadata. Both agricultural and population censuses are released at intervals of ten years by IBGE. We use data from the last three rounds of the censuses (1996, 2006 and 2017 for agricultural and 1991, 2000 and 2010 for population). The agricultural census variables of interest are: the total number and area of farms,[9] area planted with genetically modified seeds, number of tractors and machines employed in the farms, output value in the agroindustry, number of workers and total output of farms.

The population census variables: are total employment, employment shares and wages in agriculture, manufacturing and services. From the population census, we also obtain start-of-period controls, such as income per capita, the share of rural population, illiteracy rate, and population density. From Ipeadata, we obtain additional geo-climatic and socio-economic characteristics such as sectoral and total GDP, latitude, longitude, municipality area, and rural and urban population. We also use historical population census microdata from IPUMS (Min, 2019) in our auxiliary analysis (to construct Figure 3, for example).

We also use the Relacão Anual de Informações Sociais (RAIS) to obtain disaggregated

(2012).

[5]Defined as being mainly market-oriented and where "production is based on improved high yielding varieties, is fully mechanized with low labor intensity and uses optimum applications of nutrients and chemical pest, disease and weed control".

[6]In the case of cattle, we use potential yields for pasture as a proxy livestock productivity. Since extensive livestock creation composes the majority of the Brazilian production, the presence of pastures is still fundamental for cattle raising. We, therefore, assume that potential yields for pastures are a good proxy for cattle breeding productivity in our scenario.

[7]Throughout the paper, all variables representing monetary values were deflated to 2010 prices using the CPI inflation index for US dollars and the IPCA index for Brazilian Reais.

[8]For example, yerba mate and indiantea were both matched with the tea price.

[9]The two variables are reported into different intervals of area measured in hectares. We use the information to calculate the Land Gini in each municipality as described in Appendix A.2.

employment across different industries and alternative measures on total sectoral employment and employment shares in our analysis. RAIS is a yearly administrative database from the Brazilian Ministry of Labor (MTE), which provides matched employer-employee data at the individual level on the universe of formal sector employees. In our analysis, we aggregate the individual level data at the municipality level. Table 1 reports summary statistics for our main variables of interest and illustrates the evolution and difference in their values between the census years at the municipality level.

Table 1: Summary Statistics on Main Variables at Municipality Level

	2000		2010		Δ 2010-2000		
	Mean	SD	Mean	SD	Mean	SD	Observations
Panel A. Sectoral GDP							
Log Total GDP Per Capita	1.750	0.720	2.240	0.704	0.487	0.321	5,488
Log Agricultural GDP Per Capita	−0.086	1.252	0.157	1.350	0.243	0.662	5,488
Log Manufacturing GDP Per Capita	−0.430	1.180	−0.160	1.420	0.268	0.759	5,488
Log Services GDP Per Capita	1.142	0.584	1.690	0.511	0.545	0.243	5,488
Panel B. Population Census							
Employment Share Agriculture	0.422	0.203	0.354	0.182	−0.068	0.081	5,475
Employment Share Manufacturing	0.100	0.087	0.100	0.090	0.002	0.055	5,475
Employment Share Services	0.317	0.108	0.325	0.088	0.008	0.055	5,475
Log Wage in Agriculture	6.167	0.606	6.390	0.577	0.222	0.393	5,475
Log Wage in Manufacturing	6.375	0.618	6.630	0.557	0.255	0.560	5,475
Log Wage in Services	6.690	0.457	6.822	0.340	−0.300	0.422	5,475
Urban Population Share	0.590	0.233	0.641	0.219	0.052	0.063	5,475
Panel C. Exposure Measure							
Commodity Exposure	6.345	0.382	6.878	0.346	0.532	0.064	5,475
	2006		2017		Δ 2017-2006		
	Mean	SD	Mean	SD	Mean	SD	Observations
Panel D. Agricultural Census							
Log Total Farmland	10.204	1.247	10.184	1.293	−0.019	0.471	5,446
Log Machine Intensity	−5.896	1.760	−5.500	1.680	0.392	0.830	5,446
Share of Land with GE Seeds	0.019	0.086	0.080	0.193	0.060	0.171	5,446
Log Farm Productivity	−0.585	1.268	−0.481	1.260	0.103	0.830	5,446
Log Output per Worker	2.065	1.279	2.091	1.453	0.025	0.850	5,446
Land Gini	0.694	0.130	0.680	0.117	−0.018	0.084	5,446
Log Area Farms > 1.000 ha	3.834	4.926	3.051	4.720	−0.783	3.521	5,446
Log Number Farms > 1.000 ha	1.048	1.286	1.071	1.310	0.024	0.530	5,446
Log Output Value of Agroindustry	4.392	2.356	5.595	2.311	1.202	2.200	5,446

Notes: See Appendix A.1 for each variable definition and sources.

5 Empirical Strategy

The first part of the empirical analysis is to construct a shift-share (Bartik) measure to estimate the effects of the commodity price shocks on the local reallocation of production factors across the main sectors of the economy. We do so by interacting local agricultural endowment shares with the movements in global commodities prices. International prices are likely to be exogenous to local economic conditions, but local crop quantities are most likely not. Therefore, to obtain shares that are also exogenous, we combine the FAO-GAEZ potential yields with observed crop production patterns in the municipalities.

Our framework extends the settings in Dube and Vargas (2013) and Bernstein et al. (2018) by constructing exogenous agricultural shares. Following Fiszbein (2019), we construct predicted agricultural shares for each municipality by incorporating the FAO-GAEZ crop-specific attainable yields into a fractional multinomial logit (FML) model. The framework is specified as a system of equations in which the outcome variables Q_{ki} are the observed shares of each agricultural product k in total agricultural output in municipality i, and the regressors are the crop-specific potential yields vector A_i measured in tons per hectare per year.

In our baseline specification, we use the average local crop k endowment share in municipalities over the period 1995-1999 to capture the pre-period exposure to the commodity shock that started in 2000.[10] The idea behind is to create a proxy for local crop endowments shares that depends only on the FAO-GAEZ soil- and climate-based productivity measures, and therefore can be considered exogenous to local economic conditions.

The functional form of the model is given by:

$$\widehat{Q}_{ki} = E[Q_{ki,99}|A_i] = \frac{e^{\beta_k A_i}}{1 + \sum_{j=1}^{K-1} e^{\beta_j A_i}} \tag{1}$$

By construction, $\sum_{i=1}^{I} \widehat{Q}_{ki} = 1$, i.e, the predicted shares for each municipality add up to 1. The parameters are estimated by quasi-maximum-likelihood. We can motivate the FML framework by considering a simple model of optimal crop choice, as discussed in Fiszbein (2019). Assume, for instance, that farms are price-takers and choose to maximize profits. The profits obtained when choosing crop k for for a unit of farm resources l are given by the the following equation:

$$\pi_{kl} = \beta_i' A_k + \epsilon_{kl} \tag{2}$$

where π_{kl} is the profit value for each choice of crop type and farm resource and ϵ_{kl} is an error term. Then, the estimated parameters reflect the price and cost differentials among agricultural products, as well as any other factors that affect profits for different crops. If the error term is assumed to be *iid* with type I extreme value distribution, then choice k is

[10] Later, we perform robustness checks using the mean crop production over 2000-2010.

optimal (i.e. $\pi_{kl} \geq \pi_{k'l}$ for all k') with probability $\frac{e^{\beta_k A_l}}{1+\sum_{j=1}^{K-1} e^{\beta_j A_l}}$.[11]

The next step is to interact the predicted shares with the international prices for each commodity. Let P_{kt} be the international price of crop k in year t. Since a municipality can be considered a small open economy, we assume that it cannot influence, in a significant manner, the yearly movements in international commodity prices. Therefore, the price variations can be considered exogenous to local outcomes related to the structure of the economy. Our measure is then given by an annual Commodity Exposure (CE) measure for municipality i in year t:

$$CE_{it} = \sum_k \widehat{Q}_{ki} \cdot \log P_{kt} \tag{3}$$

We draw from the recent advances in the shift-share literature to guarantee that the identification assumption of the empirical design holds. In particular, our approach follows closely the assumptions in Borusyak et al. (2018), where identification is achieved via exogeneity of the shocks.[12] Our identification strategy consist of a special reduced-form case of their general shift-share instrumental variable analysis. We also apply the logarithmic transformation in the prices, and not in the whole measure, to avoid omitted variable bias as discussed in Borusyak and Hull (2020).

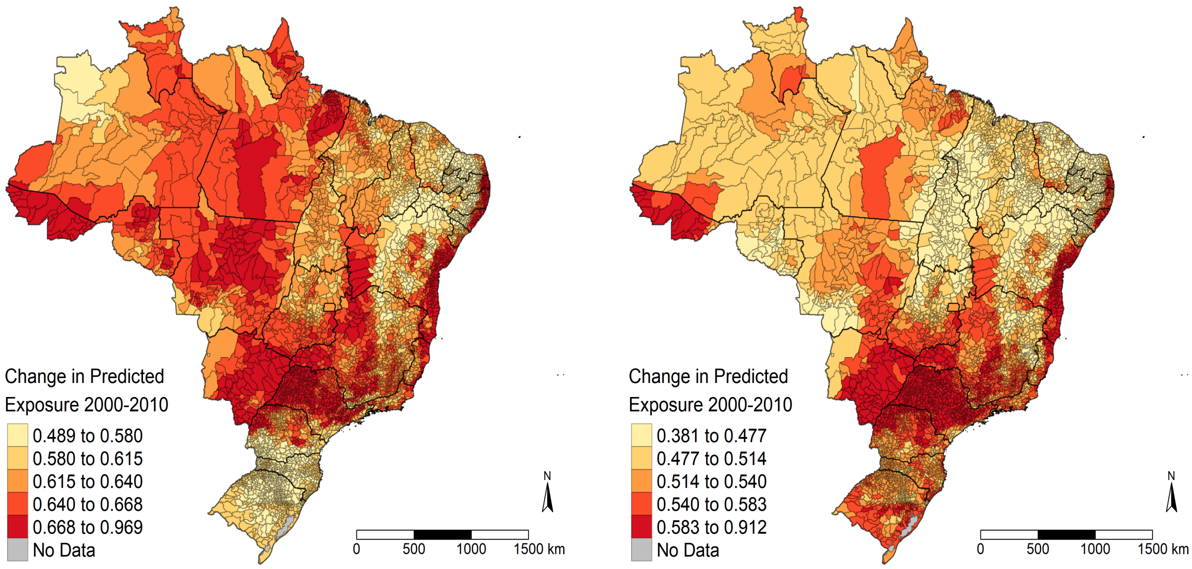
Figure 4 shows the differential exposure of each municipality to the commodity supercycle. We plot the exposure measure constructed both with and without livestock production to give a dimension of strict farming agriculture and livestock production. We show in Section 8 that our results do not change if we exclude cattle from the measure. Figure 5 shows the distribution of the values for the CE measure for both years of our analysis. The values are reasonably well distributed, giving us a large sample of exposed vs non-exposed municipalities. Figure 6 shows the change in the measure between 2000 and 2010.[13]

The exposure measure also predicts the agricultural frontier expansion that happened after 2000 in Brazil. Recall that for our baseline measure we are using pre-period crop productions combined with crop-specific attainable yields from FAO-GAEZ, which do not depend on any contemporaneous pattern of agricultural production. The measure, therefore, captures how the current agricultural frontier benefited from the commodities boom. Consider, for example, the Central-West portion of Brazil and the areas close to the Amazon Forest frontier, which are highly productive agricultural areas today and were also greatly

[11] fig. B3 shows that the predicted shares are indeed good predictors for actual (observed) shares for all crops in our sample.

[12] For a further discussion about the shift-share literature and different complementary frameworks—where identification is achieved via exogeneity of the shares—see Goldsmith-Pinkham et al. (2020).

[13] The predicted measure is also a good predictor of the actual measure (constructed with the observable crop shares). Figure B2 displays the scatter plot between the two measure values for all municipalities.



(a) Excluding Cattle

(b) Including Cattle

Figure 4: Exposure to the Commodity Shock

Notes: The maps display the spatial distribution of the change in the Commodity Exposure measure, as defined in Equation 3, between 2000 and 2010. Figure (a) displays the measure without the shares for cattle. Figure (b) displays the baseline measure that includes livestock production. See Appendix A.1 for variable definition and sources.

exposed to the shock according to our measure (Figure 4).[14]

The last step is to estimate the main equation for our analysis, which allows us to explore the effects of the commodities shocks on outcomes related to structural change in each municipality. Formally, we estimate the following equation:

$$y_{it} = \beta CE_{it} + \alpha_i + \gamma_t + \delta X_{it} + u_{it} \tag{4}$$

where y_{it} is the outcome of interest that varies across municipalities i and time t; CE_{it} is the Commodity Exposure measure as defined in Equation 3; α_i are municipality fixed effects and γ_t are time fixed effects; X_{it} is a vector of municipality control variables; u_{it} is an error term. Our identification assumption relies on the exogeneity of the CE measure in the following way:

$$\mathbb{E}\left[u_{it} | CE_{it}, X_{i,t}, \alpha_i, \gamma_t\right] = 0. \tag{5}$$

[14]The recent expansion of the agricultural frontier also generated important concerns about the possible negative environmental effects associated with the commodity boom. For a discussion, see Assunção and Bragança (2015) and Bragança (2018).

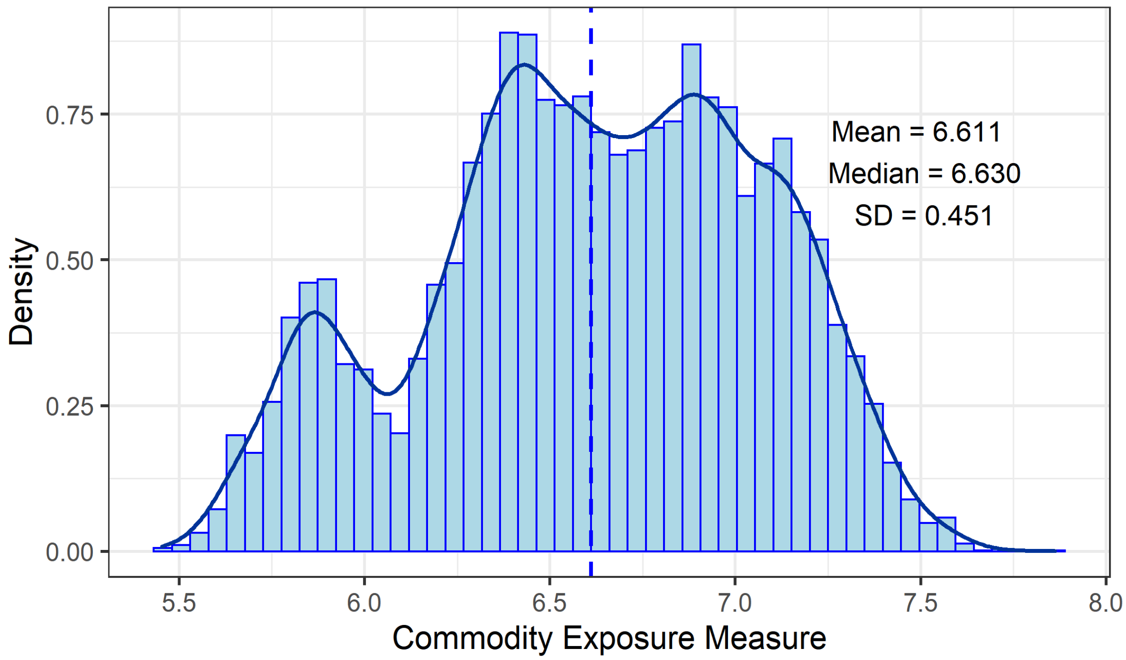


Figure 5: Histogram of the Commodity Exposure Measure

Notes: The figure displays the histogram for the Commodity Exposure measure as defined in Equation 3 for the years 2000 and 2010 together. See Appendix A.1 for variable definition and sources.

Since we constructed it using a shift-share approach, we would only need either the shares or the shocks to be exogenous in relation to the outcomes, as discussed in the approaches of Goldsmith-Pinkham et al. (2020) and Borusyak et al. (2018). Nevertheless, we argued that both the shares and shocks in our measure can be considered exogenous and, therefore, that the identification assumption holds.

In the case of sectoral employment share outcomes, our period of interest spans the ten years between the last two population censuses, which took place in 2000 and 2010. Similarly, for the agricultural censuses outcomes, we also have data on an interval of about ten years between the last two rounds (2006 and 2017). We thus estimate the first-difference version of Equation 4:

$$\Delta y_i = \beta \Delta CE_i + \delta' X_{i,1991} + \Delta u_i \tag{6}$$

where the outcome of interest Δy_i is the change in outcome variables between the last two census years and ΔCE_i is the change in the value of the exposure measure between 2000 and 2010.

We include controls for observable characteristics in the population census of 1991 to

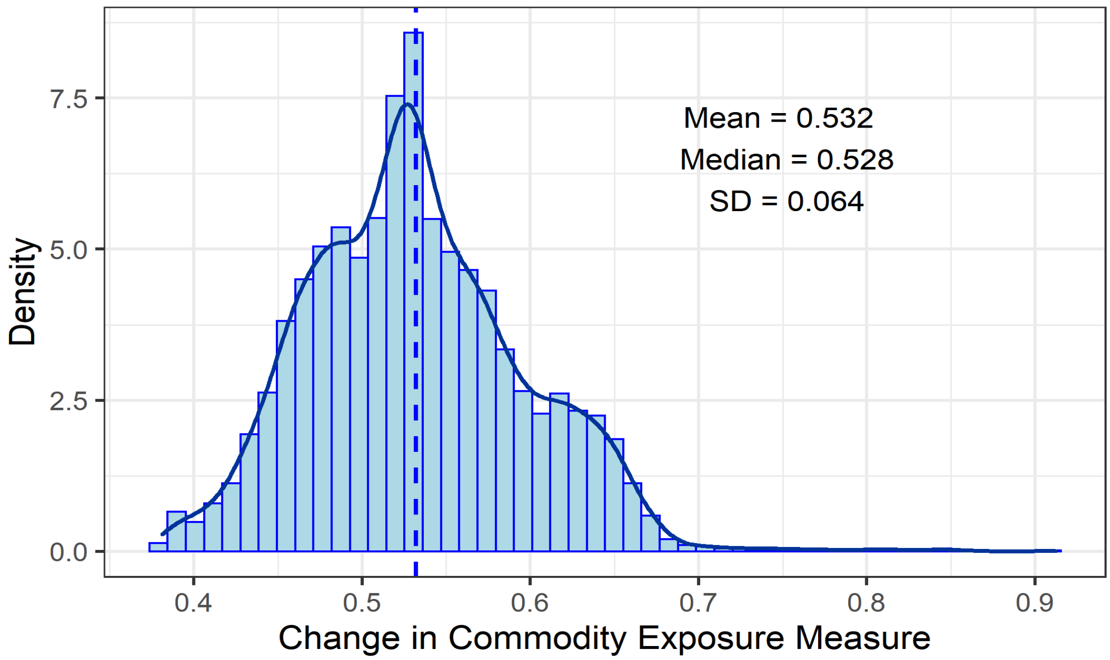


Figure 6: Histogram of the Change in the Commodity Exposure Measure between 2000-2010

Notes: The figure displays the histogram for the Commodity Exposure measure as defined in Equation 3 for the difference in the values between 2000 and 2010. See Appendix A.1 for variable definition and sources.

control for differential trends across municipalities with heterogeneous initial characteristics. In our specifications, we first add the share of rural population in 1991 and then extend the set of controls to also include population density, income per capita, illiteracy rates and region fixed effect[15] in our vector of controls. The addition of such controls is important because it can be the case that municipalities with heterogeneous initial levels of development experiences different growth paths and then our estimates could be capturing differential structural transformation trends across municipalities.[16] Table 2 displays the average values of the main observable characteristics in 1991 for municipalities that ranked below and above the median change in the exposure measure (ΔCE).[17]

We do not add any geo-climatic control to our main specification. If we controlled for any variable that was used to construct the FAO-GAEZ potential yields, the variables would

[15]Region fixed effects include dummies variables for the five macroregions of Brazil: North, Northeast, Central-West, Southeast and South.

[16]Take, for example, the state of São Paulo in the southeast region of Brazil which is one of the regions that benefited the most from the shock (as shown in Figure 4) and also one of the most developed regions in the country.

[17]Figure B1 in Appendix B complements the analysis by comparing the change in employment shares between 2000 and 2010 between the same municipalities that ranked above and below the exposure measure.

act as bad controls in our regressions. The potential yields data already incorporate climate and soil characteristics, such as rainfall, temperature, wind speed, terrain slope, and soil types, and controlling for them would, therefore, bias our estimates.[18] In Section 8.2 we perform a robustness check and add controls for other pre-period socioeconomic characteristics; geographic characteristics not incorporated in the FAO-GAEZ potential yields; and crop-specific controls.

Table 2: Comparing Municipalities Below and Above the Median Exposure to the Commodity Shock

	Below Δ CE Median	Above Δ CE Median	Difference
Agricultural Employment Share	0.570	0.430	−0.140 (0.006)
Manufacturing Employment Share	0.085	0.130	0.044 (0.003)
Share Rural Population	0.611	0.410	−0.201 (0.007)
Log Income per Capita	4.993	5.571	0.578 (0.014)
Log Population Density	2.768	3.360	0.591 (0.037)
Illiteracy Rate	0.385	0.238	−0.148 (0.004)
Observations	2,738	2,737	

Notes: See Appendix A.1 for variable definition and sources. The table presents the average values of observed characteristics of municipalities that rank below and above the median of the ΔCE measure. All variables are from the population census of 1991. The third columns reports the difference between the second and first columns. Standard errors reported in parentheses.

6 Main Results

We start our analysis by exploring how the commodity boom affected the profitability and allocation of labor across the main sectors of the Brazilian economy. As discussed in Section 1 and Section 2, the theoretical and empirical literature has shown that the expansion

[18]See Cinelli et al. (2020) for a discussion about bad controls.

of the agricultural sector can ambiguously affect the sectoral development in a economy, especially in manufacturing. It could, for instance, increase the comparative advantage of agriculture in an open economy and restrain manufacturing development. On the other hand, agglomeration and spillover effects might occur after the increase in the agricultural sector profitability. Some subsectors of the manufacturing sector could benefit from the commodity boom if their production chains depend on the production of agricultural goods. Consider, for example, industries linked to the processing of food and beverages, clothing, or even the supply of agricultural inputs, such as tractors or harvesters.

Now, we show how the commodity boom affected the profitability and growth of the major sectors. In every table, the first column (1) displays the coefficients for the regression without any controls. In the subsequent columns (2) to (4), we account for different specifications that sequentially expand the baseline set of pre-period socio-economic controls to include region fixed effects, the share of rural population, illiteracy rate, log income per capita and log population density.

Table 3 reports the estimates of the coefficient on the commodity shock exposure measure in regressions for total sectoral GDP across agriculture, manufacturing and services. The estimates in column (4), our preferred specification, imply that the commodity boom greatly benefited the agriculture sector by increasing its GDP. We also observe a first evidence on the existence of a spillover effect of the commodity shock on other sectors of the economy. The magnitudes of the coefficients imply that if a municipality experienced an increase in the CE measure equivalent to a 1 standard deviation (0.064), from the average exposure to the commodity shock, Agricultural GDP per capita would increase by about 16.5 log points (18%). Similarly, we would also observe an increase of 0.7-2.1 log points (0.7-2.2%) in Manufacturing GDP (columns (3)-(4)). The results for total GDP are positive and of about the same magnitude as the manufacturing sector. The results for the service sector GDP are negative and significant, implying a possible crowding out effect after the shock.

We now explore how labor reallocated after the commodity boom. Table 4 reports the estimates of the coefficient on the commodity shock exposure measure in regressions for employment shares in agriculture, manufacturing and services. According to our preferred specification, the estimates imply that a 1 standard deviation increase in the CE measure leads to a decrease of 0.8 percentage points (pp.) in the agricultural employment share. The estimate corresponds to about 10% of a standard deviation in the observed change of the agricultural employment share between 2000 and 2010 (8.1pp., see Table 1). In the case of manufacturing employment shares, the estimated coefficients in columns (4) indicate an increase of about 1pp. in manufacturing employment share, which corresponds to 18% of a standard deviation in the change between 2000 and 2010 (5.5pp.). We also do not observe

Table 3: The Effect of the Commodity Shock on Sectoral GDP

	(1)	(2)	(3)	(4)
Panel A.	Δ Log Total GDP per Capita			
Δ CE	−0.440***	−0.291***	0.124	0.164**
	(0.069)	(0.074)	(0.078)	(0.079)
Adj. R^2	0.007	0.013	0.045	0.050
Panel B.	Δ Log Agricultural GDP per Capita			
Δ CE	1.919***	2.007***	2.689***	2.580***
	(0.143)	(0.153)	(0.166)	(0.169)
Adj. R^2	0.034	0.034	0.069	0.075
Panel C.	Δ Log Manufacturing GDP per Capita			
Δ CE	0.516***	0.340**	0.337*	0.113
	(0.158)	(0.170)	(0.185)	(0.186)
Adj. R^2	0.002	0.003	0.043	0.062
Panel D.	Δ Log Services GDP per Capita			
Δ CE	−1.138***	−0.857***	−0.546***	−0.454***
	(0.058)	(0.060)	(0.062)	(0.060)
Adj. R^2	0.089	0.124	0.161	0.181
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,488	5,488	5,488	5,488

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2000-2010 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, C and D are 0.487 (0.321), 0.243 (0.662), 0.268 (0.759), and 0.545 (0.243). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

any significant effect on the employment share of services.[19]

We interpret the first results as providing evidence that supports the view that the commodity boom generated a significant structural transformation effect. As would be

[19] Table C1 shows similar results when considering only formal employment from RAIS. Additionally, we show in Table C2 that the effects are similar for total employment.

Table 4: The Effect of the Commodity Shock on Employment Shares

	(1)	(2)	(3)	(4)
Panel A.	Δ Employment Share in Agriculture			
Δ CE	0.053***	−0.079***	−0.095***	−0.120***
	(0.017)	(0.017)	(0.018)	(0.018)
Adj. R^2	0.002	0.071	0.072	0.086
Panel B.	Δ Employment Share in Manufacturing			
Δ CE	0.165***	0.172***	0.141***	0.151***
	(0.012)	(0.012)	(0.013)	(0.014)
Adj. R^2	0.037	0.037	0.049	0.052
Panel C.	Δ Employment Share in Services			
Δ CE	−0.152***	−0.026**	−0.013	0.001
	(0.012)	(0.012)	(0.013)	(0.013)
Adj. R^2	0.030	0.165	0.168	0.176
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2000-2010 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B and C are -0.068 (0.081), 0.002 (0.055), and 0.008 (0.055). Table C3 reports the coefficient for all the controls. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

expected, the agricultural sector benefits heavily from the commodity shock. Since our measure is multiplying prices by quantities—while also considering differentials in crop-specific agricultural productivity across locations—our results reflect how the commodity boom resulted in a significant increase in agricultural GDP. But most important than that, the boom generated a reallocation of labor away from agriculture, a result that might appear unintuitive at first.

We discuss the results in light of the recent literature that explores the localized effects of resource booms and the expansion of the agricultural sector. Bustos et al. (2016) finds that the adoption of genetically engineered (GE) soybeans across municipalities resulted in a decrease of employment share in agriculture of about 24% of a standard deviation., with an increase in employment share in manufacturing of 31% of a standard deviation. Since the GE soy is characterized as a labor-saving technology, their estimates also imply that a 1% increase in agricultural labor productivity corresponds to a 0.157pp. increase in manufacturing employment share. While our estimates are slightly lower than theirs, we also find the same sign directions after the commodity boom.

At a first moment, the reallocation of workers away from agriculture may not be expected since we are analyzing a large set of internationally traded commodities and holding crop-specific technological change constant. Moreover, many of the commodities in our sample are characterized by their labor-intensity in production and, therefore, differ fundamentally from the main hypothesis of Bustos et al. (2016). In their paper, they also look at the technical change in maize, which is found to be land-augmenting. The authors then find an opposite effect to that of the GE soy, with the expansion of maize being associated with an increase in the labor share of agriculture.[20] It could well be the case that, after the commodities boom, the agricultural sector expands its comparative advantage and draws labor in. Our results imply, however, that such events did not happen. Our main hypothesis for the mechanism that led to the observed effect is discussed in Section 7 and is related to fundamental changes in the use of capital and land in the agricultural sector.

In Table 5, we explore how the commodity super-cycle impacted urbanization and wages across the same sectors. A 1 standard deviation increase in the exposure measure led to an increase of 0.2pp. in urbanization, which corresponds to 3% of a standard deviation of the change between 2000 and 2010. Analyzing the wages outcomes, our estimates imply that a 1 standard deviation in CE is associated with an increase of about 3% in wage in the agricultural sector and an increase of 4% in services. For wages in manufacturing, we observe the opposite effect, with a negative effect of 2%. Again, our results for the manufacturing sector are similar to that of Bustos et al. (2016) for the adoption of GE soy.

Our results are also in line with the model and cross-country evidence in Gollin et al. (2016). The expansion of the agricultural sector associated with the commodity boom generated a positive urbanization effect, as predicted by their model. Although we do not observe localized effects on employment shares in the service sector, the rise in urbanization and employment share in manufacturing corroborate the "production cities" view. Moreover, our results also show that the commodity boom worked as a labor-push factor in agriculture.[21]

We now turn our attention to the possible heterogeneous effects of the commodity boom

[20] As already discussed in Section 2, we still have few evidence on the effects of agricultural technical change on structural transformation, which also appear to be context-specific (Moscona, 2018).

[21] For a discussion of push vs pull factors in the structural transformation literature, see Alvarez-Cuadrado and Poschke (2011).

Table 5: The Effect of the Commodity Shock on Urbanization and Wages

	(1)	(2)	(3)	(4)
Panel A.		Δ Urban Population Share		
Δ CE	−0.051***	0.080***	0.040**	0.032**
	(0.014)	(0.014)	(0.016)	(0.016)
Adj. R^2	0.003	0.116	0.126	0.127
Panel B.		Δ Log Wages in Agriculture		
Δ CE	0.475***	0.354***	0.307***	0.444***
	(0.080)	(0.084)	(0.092)	(0.093)
Adj. R^2	0.006	0.008	0.016	0.025
Panel C.		Δ Log Wages in Manufacturing		
Δ CE	−0.436***	−0.524***	−0.359***	−0.300**
	(0.117)	(0.124)	(0.133)	(0.136)
Adj. R^2	0.002	0.003	0.005	0.006
Panel D.		Δ Log Wages in Services		
Δ CE	1.168***	1.090***	0.572***	0.581***
	(0.085)	(0.091)	(0.094)	(0.096)
Adj. R^2	0.031	0.032	0.141	0.145
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2000-2010 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, C and D are 0.052 (0.063), 0.222 (0.393), 0.255 (0.560), and -0.300 (0.422). Table C3 reports the coefficient for all the controls in the regression. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

on the manufacturing sector. Since we observed a positive relationship between the shock and the employment share in manufacturing, we extend our analysis to test the hypothesis if the reallocation of labor was directed at a particular types of industries. To do so, we use the RAIS dataset to divide the manufacturing sector into industries that are more linked to agriculture and consider them as being part of the agroindustry. The subsectors are related to food processing, beverage production, wood processing and clothing. The other subsectors—which we label as heavy industry—are composed of chemical industries, electronics, pharmaceutical industry, metallurgy, machinery and automotive industry.[22]

One possible drawback in our analysis is the fact that RAIS only covers formal jobs, which do not fully represent the highly informalized Brazilian labor market. The manufacturing sector, nevertheless, relies intensely on workers with formal labor contracts—which is not the case in agriculture and services— and, therefore, we assume that the RAIS data approximates reasonably well the employment share inside the whole manufacturing sector. In Table 6 we show with the RAIS data that the commodity boom is associated with an increase in the manufacturing employment share of about the same magnitude that we previously obtained with the population census data.

We also obtain the output value in agroindustry from the Agricultural Census to further explore how the subsector was impacted. Our estimates imply that the production value of the agroindustry increases. After controlling for region fixed effects, the result losses significance and magnitude, but remains positive. The analysis of the employment shares outcomes shows that the reallocation of labor away from agriculture and into manufacturing appears to be equally distributed across the two subsectors that we have defined previously. A 1 standard deviation increase in the CE measure led to an increase of about 0.4pp. in agroindustry and 0.45pp. in the heavy industry.

Our results imply that the resource boom associated with the commodity supercycle generated an important reallocation of labor away from agriculture and into manufacturing and, more importantly, the expansion of the manufacturing sector is equally distributed between the agroindustry more specialized transformation industries. Such effect could imply that the structure of the economy has changed to a higher sustained growth path, if we assume that the manufacturing sector is associated with higher productivity and spillover effects, like learning-by-doing as in Matsuyama (1992) model. The scenario of an increase in the long-term growth of the economy is, however, unlikely to occur in our setting. Our analysis abstains from human capital considerations, which is a key dimension to take into consideration. Following their early work, Bustos et al. (2020b) show that the work reallocated into manufacturing after the adoption of the GE soy was relatively unskilled, which reinforced comparative advantage in the least skill-intensive manufacturing industries and slowed down local aggregate manufacturing productivity growth. We think that their effect could also be present in our analysis since agricultural labor in the Brazilian economy is characterized by low levels of human capital accumulation.

[22] See Appendix A for a detailed definition of the sector and subsector that we analyze.

Table 6: The Heterogeneous Effect of the Commodity Shock on the Type of Industry

	(1)	(2)	(3)	(4)
Panel A.	Δ Log Output in Agroindustry			
Δ CE	3.075***	2.388***	0.276	0.420
	(0.495)	(0.522)	(0.574)	(0.586)
Adj. R^2	0.008	0.010	0.040	0.040
Observations	5,447	5,447	5,447	5,447
Panel B.	Δ Employment Share in Manufacturing - RAIS			
Δ CE	0.168***	0.149***	0.102***	0.112***
	(0.025)	(0.027)	(0.029)	(0.030)
Adj. R^2	0.008	0.009	0.019	0.020
Observations	5,455	5,455	5,455	5,455
Panel C.	Δ Employment Share in Agroindustry			
Δ CE	0.101***	0.085***	0.042*	0.061***
	(0.019)	(0.020)	(0.022)	(0.023)
Adj. R^2	0.005	0.005	0.028	0.032
Observations	5,455	5,455	5,455	5,455
Panel D.	Δ Employment Share in Heavy Industry			
Δ CE	0.079***	0.070***	0.077***	0.072***
	(0.014)	(0.015)	(0.016)	(0.017)
Adj. R^2	0.006	0.006	0.011	0.011
Observations	5,455	5,455	5,455	5,455
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2006-2017 (Panel A) or 2000-2010 (Panels B, C and D) change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, C and D are 1.202 (2.200), 0.001 (0.117), -0.005 (0.093), and 0.008 (0.066). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

7 Mechanisms

We now turn to our analysis of possible channels and mechanisms that might be driving the reallocation of workers away from agriculture and into manufacturing, while also increasing the manufacturing GDP and the wages in agriculture. Our main hypothesis is based on exploring how the structure of the agriculture sector changed after the shock. We first look at the hypothesis that the commodity boom changed the use of inputs in the agriculture sector towards a more intensive use of capital in the farms. We also analyze if the commodity boom increased inequality in land ownership, increasing the land appropriated by larger farms and possibly displacing small owners from agriculture.

7.1 Use of Land, Capital Inputs, and Agricultural Productivity

In Table 7 we explore how the commodity boom impacted the use of the other two fundamental inputs used in agriculture: land and capital. Our main hypothesis in the present section is that changes in input use worked as a potential margin of adjustment to the resource boom. Our first variable of interest is the log of total farmland, to capture the expansion of land use in agricultural production. To analyze the capital adoption in farms we look at the intensity of machine use[23] and the share of agricultural land planted with genetically engineered (GE) seeds in each municipality. Our estimates in column (4) imply that an increase of 1 standard deviation in the CE measure is associated with an increase in 4.5% in machine intensity and an increase of about 1.2pp. in the share of land harvested with GE seeds. The coefficients for total farmland loses significance as we control for region fixed effects and baseline controls.

Together with our main results from the previous section, we again interpret our findings as pointing to a relevant structural transformation effect after the commodity boom. In particular, the labor that was reallocated away from agriculture and into the manufacturing sector might have been substituted by agricultural machines as our results show.[24] Moreover, the increased use of GE seeds can also explain the displacement of agricultural workers. During the last two decades, GE seeds have been extensively adopted throughout Brazil. One of the main characteristics of GE seeds is that they are resistant to herbicides and facilitate the use of no-tillage planting techniques (Bustos et al., 2016). Therefore, we would expect that an increase in the use of GE seeds leads to a decrease in the number of workers

[23]Defined as the total number of tractors, planters and harvesters in each municipality divided by the total farmland

[24]There is also a possible linkage effect that could be at play, but which we do not explore in-depth. It relates to the increase of machine intensity together with an increase in the manufacturing sector related to machine production.

per hectare needed to produce a fixed amount of agricultural output.

Our results about the absence of effects on total farmland might seem less intuitive at a first moment. Although we should expect an increase of land devoted to agricultural production in more exposed municipalities after the commodity boom, several factors might explain the lack of effect that we observe. We focus on a particular hypothesis that the landowners could have been inefficiently using their farmland and, after the shock, increased the efficiency of production without expanding their farmland devoted to agricultural production. If the hypothesis events are indeed happening, we should at least observe the farm productivity increasing.

Table 7: The Effect of the Commodity Shock on the Land Use and Capital Inputs

	(1)	(2)	(3)	(4)
Panel A.		Δ Log Total Farmland		
Δ CE	0.579***	0.500***	−0.186*	−0.111
	(0.091)	(0.097)	(0.107)	(0.108)
Adj. R^2	0.006	0.007	0.052	0.062
Panel B.		Δ Log Machine Intensity		
Δ CE	0.264	0.547**	0.582**	0.680***
	(0.206)	(0.217)	(0.234)	(0.237)
Adj. R^2	0.000	0.003	0.020	0.023
Panel C.		Δ Share of Land with GE Seeds		
Δ CE	0.390***	0.379***	0.188***	0.181***
	(0.031)	(0.032)	(0.028)	(0.029)
Adj. R^2	0.021	0.021	0.191	0.193
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,446	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2006-2017 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B and C are -0.019 (0.471),0.392 (0.830), and 0.060 (0.171). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

In Table 8 we show that it appears to be the case. We measure farm productivity as the total output of agriculture divided by the total farmland in hectares. Our results show that the commodity boom is positively associated with farm productivity. The results also represent an important mechanism that might be driving our main results regarding structural transformation. The evidence on farmland productivity corroborates our findings in Section 6 about the commodity boom working as a labor-push factor in agriculture.

We also explore if labor productivity in the agricultural sector increased after the commodities boom. If we observed a positive and significant effect of the exposure to the commodity boom on labor productivity, it could mean that the effects of labor-saving vs labor-augmenting technologies are missing from our analysis. The prevalence of one type of technological change over another might be important in explaining the structural transformation patterns we observe, as discussed in Bustos et al. (2016). We do not observe, however, any statistically significant effect in our preferred specification. Since the literature has well-documented evidence on the adoption of both types of agricultural technical change in Brazil during the period which we analyze, it could be the case that their overall effect over structural transformation eventually cancels out.

7.2 Inequality in Land Ownership

We now explore how the commodity boom impacted the reallocation of farmland and land inequality across municipalities and how they might have worked as an important mechanism in driving the structural change effects we documented in the previous sections. Land inequality has been more extensively studied in the broad growth literature, but it is often overlooked in the structural transformation literature. One of the main hypothesis linked to the study of land inequality is related to the works of Engerman and Sokoloff (1994) and Galor et al. (2009), which suggest that inequality in the distribution of landownership had adverse effects on the emergence of public schooling during the transition from an agricultural economy to an industrial one. The hypothesis has been extensively tested in historical settings, implying that land inequality indeed played an important role in respect to long term economic development (Easterly, 2007; Vollrath, 2007; Cinnirella and Hornung, 2016; Wigton-Jones, 2020).

In our setting, the results regarding land inequality, institutions and human capital are unlikely to have a significant effect, mainly because we are analyzing a modern economy that has already experienced a major industrialization and expansion of public education. Nevertheless, we still propose that land inequality might be playing an important role in the relationship between commodity booms and structural transformation, in particular via

Table 8: The Effect of the Commodity Shock on Agricultural Productivity

	(1)	(2)	(3)	(4)
Panel A.		Δ Log Farm Productivity		
Δ CE	1.012***	1.009***	0.399**	0.567***
	(0.180)	(0.196)	(0.193)	(0.194)
Adj. R^2	0.006	0.006	0.096	0.118
Panel B.		Δ Log Output per Worker		
Δ CE	1.324***	1.434***	0.033	0.215
	(0.191)	(0.207)	(0.197)	(0.197)
Adj. R^2	0.010	0.010	0.124	0.144
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,446	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2006-2017 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A and B are 0.103 (0.830), and 0.025 (0.850). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

productivity changes and displacement of small landowners.

The relation between land inequality and agricultural productivity has been studied in Vollrath (2007). In a cross-sectional study of countries, the author finds that the Gini coefficient for land holdings is negatively associated with productivity in farms. In our setting, the relation between the commodities boom, land inequality and structural change could go plausibly in different directions. We expect that, after the commodity shock, land inequality increases given the structure of the Brazilian agricultural sector and its high and persistent inequality in landholdings discussed in Section 3. The relationship between land inequality and agricultural productivity would then be ambiguous a priori. We could expect that productivity decreases given the rise in land inequality, supporting Vollrath's findings. On the other hand, productivity could also be positively related to productivity in our case. Since commodities such as soy, tobacco, cotton, sugarcane and cattle raising are associated with large and productive farms, it could be the case that, after the shock, both land inequality and farm productivity increase in Brazil.[25]

In Table 9 we explore how the commodity boom affected the Land Gini taking short and long differences. We choose to report both types of results because land inequality, as a particular form of wealth inequality, is very persistent and changes slowly over time. Our results imply that an increase of 1 standard deviation in the exposure measure led to an increase of 0.004 points in the short difference Land Gini (between the 2006 and 2017 Agricultural Censuses) and an increase of about 0.007 points in the long difference (between 1995 and 2017). The two effects account for about 5% of a standard deviation in the observed change.

We complement the pieces of evidence by exploring in Table 10 the number of farms with more than 1000 hectares and the area appropriated by them in each municipality. We choose such arbitrary cutoff of total area because it captures the average farm size of the top 1% of agricultural establishments with respect to area appropriation, as reported in the Agricultural Census and discussed in Section 3. We find that an increase of 1 standard deviation in the exposure measure is associated with an increase of 0.42% in the area of the top 1% farms in the short difference. The coefficient associated with the long difference is not statistically significant, although higher in magnitude. Analyzing the number of farms with more than 1000 hectares, we find that an increase of 1 standard deviation in the CE measure leads to an increase of 2.1pp. in the short difference and an increase of 8pp. in the long difference.

Together with the evidence in Table 8, it appears to be the case that both agricultural productivity and land inequality increased, supporting the view that farms associated with large field crops increased their productivity after the commodity shock. Moreover, since we observe both the Land Gini and the area and number of top 1% farms increasing without the same happening with total farmland, we interpret the results as also supporting the view that the commodity shock impacted structural transformation via the displacement of small landowners and workers away from agriculture.

8 Robustness Checks

8.1 Alternative Standard Errors

Figure 4 suggests that the Commodity Exposure measure is correlated across space. We then consider a series of possible specifications of alternative standard errors to account for spatial

[25] Although smaller farms were usually associated with higher productivity in the empirical literature, there is a recent divergence between the microeconomic and macroeconomic evidence. For a discussion, see Sanchez et al. (2019).

Table 9: The Effect of the Commodity Shock on Land Inequality

	(1)	(2)	(3)	(4)
Panel A.		Δ Land Gini (2006-2017)		
Δ CE	0.179***	0.163***	0.072***	0.065***
	(0.017)	(0.018)	(0.021)	(0.021)
Adj. R^2	0.018	0.019	0.065	0.071
Observations	5,446	5,446	5,446	5,446
Panel B.		Δ Land Gini (1995-2017)		
Δ CE	0.217***	0.229***	0.127***	0.111***
	(0.022)	(0.023)	(0.025)	(0.025)
Adj. R^2	0.019	0.019	0.106	0.111
Observations	4,919	4,919	4,919	4,919
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2006-2017 or 1995-2017 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A and B are -0.018 (0.084), and -0.007 (0.100). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

correlation. First, we cluster the standard errors at a higher unit of observation to allow the residuals to be correlated within geographical areas larger than a single municipality. To do so, we cluster at the microregion level, which roughly represents Brazilian local labor market regions (Adão, 2015). The approach accounts for residual auto-correlation and spatial spillovers across nearby municipalities with economic links. We also cluster at a even higher unit of observation, the mesoregion, defined by IBGE, which are larger than microregions but smaller than the five macroregions.[26] Second, we calculate standard errors that correct for spatial dependence using different distance cutoffs as suggested by Conley (1999).[27]

Our shift-share design can also suffer from a particular problem of inference, that lead to

[26] In 2010, there were 558 microregions and 137 mesoregions in Brazil.

[27] We implement the Conley standard errors using *acreg* command created by Colella et al. (2019) in Stata.

Table 10: The Effect of the Commodity Shock on Land Inequality

	(1)	(2)	(3)	(4)
Panel A.	Δ Log Area Farms > 1.000 ha (2006-2017)			
Δ CE	0.179***	0.163***	0.072***	0.065***
	(0.017)	(0.018)	(0.021)	(0.021)
Adj. R^2	0.018	0.019	0.065	0.071
Panel B.	Δ Log Area Farms > 1.000 ha (1995-2017)			
Δ CE	1.267	6.090***	2.087*	1.544
	(0.970)	(1.022)	(1.094)	(1.114)
Adj. R^2	0.000	0.030	0.086	0.094
Panel C.	Δ Log Number Farms > 1.000 ha (2006-2017)			
Δ CE	0.848***	0.827***	0.294**	0.336**
	(0.117)	(0.125)	(0.134)	(0.136)
Adj. R^2	0.010	0.010	0.031	0.033
Panel D.	Δ Log Number Farms > 1.000 ha (1995-2017)			
Δ CE	1.009***	1.839***	1.323***	1.228***
	(0.148)	(0.156)	(0.165)	(0.169)
Adj. R^2	0.008	0.043	0.073	0.079
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,446	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2006-2017 or 1995-2017 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, C and D are -0.783 (3.522), -2.790 (4.573), 0.024 (0.530), and 0.030 (0.722). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

a overrejection problem. Regression residuals can be correlated across regions with similar shares, independently of their geographic location, and the OLS true variability may then be understated when using the usual clustering or heteroskedasticity-robust errors. We then apply the AKM and AKM0 corrections proposed by Adao et al. (2019) to account for the dependence between residuals generated by unobserved shift-share components.

Tables 11, 12 and 13 report the alternative standard errors for our main results regarding the effect of the commodity on structural change, input use and land inequality. We report coefficients for our main specification (column (4) in the previous tables), as well as the coefficients for the main specification but without region fixed effects. We remove the region fixed effects because the *reg_ss* package developed by Adao et al. (2019) that implements the AKM and AKM0 corrections does not allow for fixed effects regressions.

The significance of our results remain the same in respect to employment shares and wages. Only the effects on urban population share appear to loose significance when we account for spatial correlation. The significance of the coefficients for input use also remain mostly the same, with only the coefficient for spatial correlation above 200km losing significance. The results for land inequality also remain statistically significant with the exception of the coefficient for the short difference in the Land Gini for the 400km cutoff.

8.2 Additional Controls

A potential concern regarding our analysis is that specific commodities, such as soy, meat or sugarcane, are driven the results. Moreover, municipalities more exposed to the commodities boom can also have higher overall agricultural productivity or different labor force composition prior to the shock. Our estimates could then be capturing specific specialization in commodity production or differential trends of structural transformation across locations. To address the differential trends, we report estimates of Equation 6 including the following additional controls: share of women in the labor force in 1991; land productivity in 1995 (measured as the log value of output per total farmland); a dummy variable if a commodity represents individually more than 30% of a municipality's total agricultural production; and a dummy variable if a group of commodities (meat, coffee, maize, soybean and sugarcane) represents together more than 50% a municipality's total agricultural production. We also add a control for altitude—which is an important geographical characteristic not incorporated in the FAO-GAEZ potential yields model—and other proxies for access to markets, such as distance to the federal capital; distance to the state capital; and state capital dummy.

Tables C4 and C5 report the coefficients for the main regressions with the full set of additional controls. The sign of estimated coefficients remains the same, while also remaining significant and of similar magnitude for the main outcomes (the only exception is the coefficient for urbanization). Overall, our results remain stable over the inclusion of a set of additional variables that controls for initial agricultural development, labor force composition, geographical characteristics, and prevalence of specific commodities.

8.3 Alternative Measures

We now consider alternative definitions for our Commodity Exposure Measure. Based on the continuous measure defined in Equation 3, we construct dummy variables that equal one if the difference in the CE measure is in some given top or bottom percentile of its distribution. We could then interpret that a municipality is "treated" if the dummy variable equals one. For the percentile thresholds, we construct dummy variables for the top and bottom 10th and 25th percentiles, as well as for the values above the median of the distribution.

Tables 14, 15 and 16 report the coefficients for our preferred specification on the main outcomes using the alternative measures. Overall, the results remain qualitatively the same, but the interpretation of the coefficients is different. Now, we are comparing the municipalities above or below an arbitrary threshold of the CE measure distribution with the rest of the municipalities. In respect to employment shares, the results show that if a municipality is the the top 10% or top 25% of the Δ CE measure distribution, it experiences a decrease in employment share in agriculture of 2.1-2.4 pp. and an increase of 2.5-2.7pp. in employment share in manufacturing. The bottom dummies also corroborate our main findings. Being in the bottom distribution of the Δ CE measure is associated with no effect on employment shares, urbanization, and wages. In some cases there are even opposite effects in comparison with the municipalities in the top of the distribution. The results on input use and land inequality also remain qualitatively the same except for the machine intensity variable.

We also perform robustness check regarding the definition of the Commodity Exposure measure constructed in Equation 3. First, we use the mean commodity production in PAM for actual the supercycle period (2000-2010) to construct the shares in the FML model. Second, we exclude cattle shares from the baseline measure. Third, we use the FAO-GAEZ high inputs in the FML model. Tables C6, C7 and C8 report the estimates for the three different definitions for the CE measure. The results remain consistent with the main estimates in both magnitude and significance.

8.4 Observations at a Large Aggregation Level

A potential issue in the empirical analysis performed so far is that some municipalities might be to small to capture flows between rural and urban areas, which are usually associated with agricultural and manufacturing labor. We then aggregate our municipality-level observations at the microregion level, which represents a larger unit of observation and, as discussed previously, also represents Brazilian local labor markets. Tables C9, C10 and C11 report the baseline estimates at the microregion-level. The results have similar magnitude and are consistent with the estimates reported in Section 6.

Table 11: Alternative Standard Errors for Population Census Outcomes

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Δ CE - Main Specification	−0.120***	0.151***	0.001	0.0325**	0.444***	−0.300**	0.581***
	(0.018)	(0.014)	(0.013)	(0.016)	(0.093)	(0.136)	(0.096)
Adj. R^2	0.086	0.052	0.176	0.127	0.025	0.006	0.145
Cluster at microregion	(0.033)***	(0.033)***	(0.015)	(0.021)	(0.136)***	(0.141)**	(0.182)***
Cluster at mesoregion	(0.046)**	(0.047)***	(0.019)	(0.025)	(0.162)***	(0.161)*	(0.209)***
Conley 50km	(0.036)***	(0.031)***	(0.018)	(0.023)	(0.137)***	(0.165)*	(0.173)***
Conley 100km	(0.049)**	(0.044)***	(0.022)	(0.028)	(0.173)**	(0.167)*	(0.236)**
Conley 200km	(0.065)*	(0.061)**	(0.024)	(0.032)	(0.186)**	(0.157)*	(0.267)**
Conley 400km	(0.071)*	(0.073)**	(0.020)	(0.014)**	(0.141)***	(0.194)	(0.171)***
Δ CE - No Region FE	−0.138***	0.174***	0.006	0.045***	0.524***	−0.345***	0.575***
	(0.018)	(0.014)	(0.012)	(0.016)	(0.089)	(0.131)	(0.094)
Adj. R^2	0.082	0.037	0.171	0.123	0.014	0.005	0.100
AKM	(0.035)***	(0.044)***	(0.021)	(0.015)***	(0.121)***	(0.086)***	(0.224)**
AKM0	(0.055)**	(0.070)**	(0.033)	(0.027)	(0.189)**	(0.143)*	(0.353)*
Observations	5,475	5,475	5,475	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents the main specification estimates of Equation 6 as reported in column (4) of Tables 4 and 5. Conley (1999) standard errors are computed using the *acreg* command created by Colella et al. (2019). AKM and AKM0 corrections are computed using the *reg_ss* command created by Adao et al. (2019). Estimates of the main specification without region fixed effects are also reported since *reg_ss* does not allow for fixed effects. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 12: Alternative Standard Errors for Input Use

	Δ Log Total Farmland	Δ Log Machine Intensity	Δ Share of Land with GE Seeds
Δ CE - Main Specification	−0.111	0.680***	0.181***
	(0.108)	(0.237)	(0.029)
Adj. R^2	0.062	0.023	0.193
Cluster at microregion	(0.168)	(0.295)**	(0.073)**
Cluster at mesoregion	(0.211)	(0.340)**	(0.084)**
Conley 50km	(0.153)	(0.313)**	(0.074)**
Conley 100km	(0.194)	(0.331)**	(0.093)*
Conley 200km	(0.254)	(0.466)	(0.084)**
Conley 400km	(0.300)	(0.624)	(0.087)**
Δ CE - No Region FE	0.238**	0.597***	0.110***
	(0.102)	(0.226)	(0.030)
Adj. R^2	0.042	0.006	0.082
AKM	(0.144)***	(0.298)**	(0.155)
AKM0	(0.187)	(0.518)**	(0.246)
Observations	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents the main specification estimates of Equation 6 as reported in column (4) of Table 7. Conley (1999) standard errors are computed using the *acreg* command created by Colella et al. (2019). AKM and AKM0 corrections are computed using the *reg_ss* command created by Adao et al. (2019). Estimates of the main specification without region fixed effects are also reported since *reg_ss* does not allow for fixed effects. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 13: Alternative Standard Errors for Land Inequality

	Δ Land Gini (2006-2017)	Δ Land Gini (1995-2017)
Δ CE - Main Specification	0.065***	0.111***
	(0.021)	(0.025)
Adj. R^2	0.071	0.111
Cluster at microregion	(0.031)**	(0.042)***
Cluster at mesoregion	(0.033)*	(0.051)**
Conley 50km	(0.027)**	(0.044)**
Conley 100km	(0.033)*	(0.056)**
Conley 200km	(0.032)*	(0.059)*
Conley 400km	(0.039)	(0.062)*
Δ CE - No Region FE	0.062***	0.108***
	(0.020)	(0.024)
Adj. R^2	0.067	0.083
AKM	(0.030)**	(0.048)**
AKM0	(0.053)	(0.078)
Observations	5,446	4,919

Notes: See Appendix A.1 for variable definition and sources. The table presents the main specification estimates of Equation 6 as reported in column (4) of Table 9. Conley (1999) standard errors are computed using the *acreg* command created by Colella et al. (2019). AKM and AKM0 corrections are computed using the *reg_ss* command created by Adao et al. (2019). Estimates of the main specification without region fixed effects are also reported since *reg_ss* does not allow for fixed effects. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%

Table 14: The Effect of the Commodity Shock on Employment Shares - Alternative Measures

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Top 10% Δ CE	−0.024***	0.027***	−0.002	0.005**	0.077***	−0.007	0.077***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.017)	(0.023)	(0.016)
Adj. R^2	0.088	0.052	0.176	0.127	0.025	0.005	0.142
Top 25% Δ CE	−0.021***	0.025***	−0.002	0.003	0.074***	−0.018	0.098***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.013)	(0.019)	(0.013)
Adj. R^2	0.091	0.061	0.176	0.127	0.027	0.005	0.147
Above Median Δ CE	0.0002	0.009***	−0.005***	0.003*	0.033***	−0.035**	0.077***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.012)	(0.017)	(0.013)
Adj. R^2	0.080	0.036	0.177	0.127	0.023	0.006	0.145
Bottom 10% Δ CE	0.007*	0.000	−0.008***	0.003	−0.011	0.019	0.0303
	(0.004)	(0.002)	(0.003)	(0.003)	(0.019)	(0.029)	(0.020)
Adj. R^2	0.081	0.031	0.178	0.126	0.022	0.005	0.140
Bottom 25% Δ CE	0.001	−0.006***	0.000	−0.004*	0.012	0.054**	0.012
	(0.003)	(0.002)	(0.002)	(0.003)	(0.015)	(0.023)	(0.016)
Adj. R^2	0.080	0.033	0.176	0.127	0.022	0.006	0.139
Observations	5,475	5,475	5,475	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Tables 4 and 5 substituting the continuous Δ CE measure with dummy variables that indicate if a municipality is at the top and bottom 10th and 25th percentiles and above the median in the distribution of the Δ CE measure. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table 15: The Effect of the Commodity Shock on the Land Use and Capital Inputs - Alternative Measures

	Δ Log Total Farmland	Δ Log Machine Intensity	Δ Share of Land with GE Seeds
Top 10% Δ CE	−0.0240	−0.093***	0.001
	(0.019)	(0.029)	(0.007)
Adj. R^2	0.062	0.022	0.190
Top 25% Δ CE	−0.027*	−0.026	0.028***
	(0.015)	(0.023)	(0.007)
Adj. R^2	0.063	0.021	0.194
Above Median Δ CE	−0.004	0.040*	0.021***
	(0.016)	(0.025)	(0.004)
Adj. R^2	0.062	0.022	0.193
Bottom 10% Δ CE	0.012	−0.309***	−0.004**
	(0.022)	(0.058)	(0.001)
Adj. R^2	0.062	0.032	0.190
Bottom 25% Δ CE	0.003	−0.209***	−0.008***
	(0.021)	(0.039)	(0.002)
Adj. R^2	0.062	0.029	0.191
Observations	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 7 substituting the continuous Δ CE measure with dummy variables that indicate if a municipality is at the top and bottom 10th and 25th percentiles and above the median in the distribution of the Δ CE measure. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table 16: The Effect of the Commodity Shock on Land Inequality - Alternative Measures

	Δ Land Gini (2006-2017)	Δ Land Gini (1995-2017)
Top 10% Δ CE	0.014***	0.011***
	(0.004)	(0.004)
Adj. R^2	0.075	0.108
Top 25% Δ CE	0.009***	0.01***
	(0.003)	(0.003)
Adj. R^2	0.074	0.108
Above Median Δ CE	0.005*	0.009***
	(0.003)	(0.003)
Adj. R^2	0.073	0.108
Bottom 10% Δ CE	0.008**	−0.008*
	(0.004)	(0.005)
Adj. R^2	0.073	0.108
Bottom 25% Δ CE	0.001	−0.013***
	(0.003)	(0.004)
Adj. R^2	0.072	0.109
Observations	5,446	4,919

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 9 substituting the continuous Δ CE measure with dummy variables that indicate if a municipality is at the top and bottom 10th and 25th percentiles and above the median in the distribution of the Δ CE measure. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

9 Conclusion

We show that the 2000s commodity supercycle led to a significant structural change in Brazilian municipalities between 2000 and 2010. Our identification strategy relies on a shift-share measure that is constructed interacting climate- and soil-based measures of crop-specific potential yields with the international prices of commodities. The measure gives us cross-sectional and time variation of local exposure to the commodity shock. We find that, after the boom, labor was reallocated away from agriculture and towards the manufacturing sector. We also explored the possible mechanisms that are driving such effect and find that substitution of labor by capital in the farms and increasing inequality among landowners are important and previously unexplored factors that might be driving the observed effects.

We also interpret our results in light of the broad Dutch Disease literature and the Prebisch-Singer hypothesis. The resource boom did not appear to crowd out the manufacturing sector, which shows that the usual resource curse effects were not present in the context analyzed. If we take into account the evidence presented in present paper together with that in Bustos et al. (2016), it appears that changes in the structure of the modern agricultural sector in Brazil—via exogenous prices variations or technological change—can indeed generate structural transformation.

Although deindustrialization and low trends of growth and productivity are indeed salient features in the Brazilian economy during the last decades, our work provides evidence that the commodity boom is not one of the fundamental causes of such issues. We raise the hypothesis that the usual suspects, such as the overall quality of institutions and the accumulation and quality of human capital might be the key factor preventing the betterment of productivity and economic growth in the economy. We do not, however, test directly for the human capital hypothesis and leave it for future work.

References

Adão, Rodrigo, "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from Brazil," 2015. Working Paper.

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Allcott, Hunt and Daniel Keniston, "Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America," *The Review of Economic Studies*, 2018, *85* (2), 695–731.

Alvarez-Cuadrado, Francisco and Markus Poschke, "Structural change out of agriculture: Labor push versus labor pull," *American Economic Journal: Macroeconomics*, 2011, *3* (3), 127–58.

Assunção, Assunção and Arthur Bragança, "Technological change and deforestation: evidence from the Brazilian soybean revolution," 2015. Working Paper.

Bartik, Timothy J, "Who benefits from state and local economic development policies?," 1991.

Bauluz, Luis, Yajna Govind, and Filip Novokmet, "Global Land Inequality," 2020. Working Paper.

Benguria, Felipe, Felipe Saffie, and Sergio Urzúa, "The transmission of commodity price super-cycles," 2018. Working Paper.

Bernstein, Shai, Emanuele Colonnelli, Davide Malacrino, and Timothy McQuade, "Who creates new firms when local opportunities arise?," 2018. Working Paper.

Blanchard, Olivier Jean, Lawrence F Katz, Robert E Hall, and Barry Eichengreen, "Regional evolutions," *Brookings papers on economic activity*, 1992, *1992* (1), 1–75.

Borusyak, Kirill and Peter Hull, "Non-Random Exposure to Exogenous Shocks: Theory and Applications," 2020. Working Paper.

_ , _ , and Xavier Jaravel, "Quasi-experimental shift-share research designs," 2018. Working Paper.

Bragança, Arthur, "The economic consequences of the agricultural expansion in Matopiba," *Revista Brasileira de Economia*, 2018, *72* (2), 161–185.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural productivity and structural transformation: Evidence from Brazil," *American Economic Review*, 2016, *106* (6), 1320–65.

_ **, Gabriel Garber, and Jacopo Ponticelli**, "Capital accumulation and structural transformation," *The Quarterly Journal of Economics*, 2020, *135* (2), 1037–1094.

_ **, Juan Manuel Castro-Vincenzi, Joan Monras, and Jacopo Ponticelli**, "Industrialization without Innovation," 2020. Working Paper.

Cavalcanti, Tiago, Daniel Da Mata, and Frederik Toscani, "Winning the oil lottery: The impact of natural resource extraction on growth," *Journal of Economic Growth*, 2019, *24* (1), 79–115.

Cinelli, Carlos, Andrew Forney, and Judea Pearl, "A Crash Course in Good and Bad Controls," *Available at SSRN 3689437*, 2020.

Cinnirella, Francesco and Erik Hornung, "Landownership concentration and the expansion of education," *Journal of Development Economics*, 2016, *121*, 135–152.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with arbitrary clustering," 2019.

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of econometrics*, 1999, *92* (1), 1–45.

Corden, W Max and J Peter Neary, "Booming sector and de-industrialisation in a small open economy," *The economic journal*, 1982, *92* (368), 825–848.

Costa, Francisco, Jason Garred, and Joao Paulo Pessoa, "Winners and losers from a commodities-for-manufactures trade boom," *Journal of International Economics*, 2016, *102*, 50–69.

Costinot, Arnaud, Dave Donaldson, and Cory Smith, "Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world," *Journal of Political Economy*, 2016, *124* (1), 205–248.

der Ploeg, Frederick Van, "Natural resources: curse or blessing?," *Journal of Economic literature*, 2011, *49* (2), 366–420.

Dube, Oeindrila and Juan F Vargas, "Commodity price shocks and civil conflict: Evidence from Colombia," *The review of economic studies*, 2013, *80* (4), 1384–1421.

Easterly, William, "Inequality does cause underdevelopment: Insights from a new instrument," *Journal of development economics*, 2007, *84* (2), 755–776.

Engerman, Stanley L and Kenneth L Sokoloff, "Factor endowments: institutions, and differential paths of growth among new world economies: a view from economic historians of the United States," *NBER Working Paper*, 1994.

Erten, Bilge and José Antonio Ocampo, "Super cycles of commodity prices since the mid-nineteenth century," *World development*, 2013, *44*, 14–30.

FGV, "Agribusiness in Brazil: an overview," *FGV Projetos*, 2015.

Fiszbein, Martin, "Agricultural diversity, structural change and long-run development: Evidence from the US," 2019. Working Paper.

Galor, Oded, Omer Moav, and Dietrich Vollrath, "Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence," *The Review of economic studies*, 2009, *76* (1), 143–179.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gollin, Douglas, Remi Jedwab, and Dietrich Vollrath, "Urbanization with and without industrialization," *Journal of Economic Growth*, 2016, *21* (1), 35–70.

_ **, Stephen Parente, and Richard Rogerson**, "The role of agriculture in development," *American economic review*, 2002, *92* (2), 160–164.

Hansen, Gary D. and Edward C. Prescott, "Malthus to Solow," *American Economic Review*, 2002, *92* (4), 1205–1217.

Harvey, David I, Neil M Kellard, Jakob B Madsen, and Mark E Wohar, "The Prebisch-Singer hypothesis: four centuries of evidence," *The review of Economics and Statistics*, 2010, *92* (2), 367–377.

IIASA/FAO, "Global Agro-ecological Zones (GAEZ v3.0)," 2012.

Lewis, Arthur R., "Economic development with unlimited supply of labor," *Manchester School of Economic and Social Studies*, 1954, *22*, 139—-191.

Matsuyama, Kiminori, "Agricultural productivity, comparative advantage, and economic growth," *Journal of economic theory*, 1992, *58* (2), 317–334.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik, "Institutions and the resource curse," *The economic journal*, 2006, *116* (508), 1–20.

Minnesota Population Center, *Integrated Public Use Microdata Series, International: Version 7.2* 2019.

Moscona, J, "Agricultural development and structural change within and across countries," 2018. Working Paper.

Ngai, L Rachel and Christopher A Pissarides, "Structural change in a multisector model of growth," *American economic review*, 2007, *97* (1), 429–443.

Nunn, Nathan, "Slavery, inequality, and economic development in the americas," *Institutions and economic performance*, 2008, *15*, 148–180.

Ploeg, Frederick Van Der and Steven Poelhekke, "The impact of natural resources: Survey of recent quantitative evidence," *The Journal of Development Studies*, 2017, *53* (2), 205–216.

PNUD, IPEA, "Atlas do Desenvolvimento Humano no Brasil," 2020.

Prebisch, Raúl, "The Economic Development of Latin America and Its Principal Problems," 1950.

PwC, "Agribusiness in Brazil: an overview," 2013.

Rostow, Walt Whitman, *The stages of economic growth: A non-communist manifesto*, Cambridge university press, 1960.

Sachs, Jeffrey D and Andrew M Warner, "The big push, natural resource booms and growth," *Journal of development economics*, 1999, *59* (1), 43–76.

_ **and** _ , "The curse of natural resources," *European economic review*, 2001, *45* (4-6), 827–838.

Sanchez, Fernando M Aragon, Diego Restuccia, and Juan Pablo Rud, "Are small farms really more productive than large farms?," 2019. Working Paper.

Singer, HW, "The Distribution of Gains between Investing and Borrowing Countries," *The American Economic Review*, 1950, *40* (2), 473–485.

Uribe-Castro, M, "Caffeinated development: exports, human capital, and structural transformation in Colombia," *Working Paper*, 2019, *22.*

Vollrath, Dietrich, "Land distribution and international agricultural productivity," *American Journal of Agricultural Economics*, 2007, *89* (1), 202–216.

_ , "The agricultural basis of comparative development," *Journal of Economic Growth*, 2011, *16* (4), 343–370.

Wigton-Jones, Evan, "Legacies of inequality: the case of Brazil," *Journal of Economic Growth*, 2020, *25* (4), 455–501.

Appendix

A Data

The present section contains a detailed description and the sources of the main variables used in the the paper.

A.1 Definitions and Sources

Variables for Constructing the Commodity Exposure Measure

Observed shares of each crop ($Q_{ki}, 99$): Data on the observed shares of each crop between 1995 and 1999 is from the Produção Agrícola Municipal (PAM) and Pesquisa da Pecuária Municipal (PPM) and it is sourced from the IBGE online data repository SIDRA. We construct the variable by dividing the quantities of each crop in each municipality-year observation by the total quantity of all crops used in our sample for each observation. Table A1 reports the number of municipalities producing each type of crop. Data for the quantities of seasonal and permanent crops in PAM is from Tabela 5457. Data for the number of cattle heads in PPM if from Tabela 3939.

Crop-specific potential yields: Data on potential yields comes from the Global Agro-Ecological Zones project v3.0 (IIASA/FAO (2012)), which is publicly available on their website.[28] The variable measures agricultural suitability as the maximum attainable yields for a crop in a certain geographical area and are reported in tons per hectare per year for each crop in each grid cell of 0.083x0.083 degrees. FAO uses information on climatic conditions—including precipitation, temperature, wind speed, sunshine hours and relative humidity— together with data on soil, topography and elevation to determine the maximum attainable yields. FAO's data also enable us to choose different types of inputs and water supply conditions. We use data on intermediate levels of inputs/technology and rain-fed conditions, which better approximates our scenario.

The definition of each inputs/technology in the FAO-GAEZ dataset documentation is as follows. Low-level inputs/traditional management: "Under a low level of inputs (traditional management assumption), the farming system is largely subsistence based. Production is based on the use of traditional cultivars (if improved cultivars are used, they are treated in the same way as local cultivars), labor intensive techniques, and no application of nutrients, no use of chemicals for pest and disease control and minimum conservation mea-

[28] http://www.gaez.iiasa.ac.at/.

Table A1: Agricultural Endowments Across Municipalities

	Number of Municipalities	
	2000	2010
Cattle	5471	5518
Maize	5329	5176
Rice	4071	3084
Banana	3795	3555
Orange	3634	3006
Sugarcane	3483	3695
Coffee	2008	1822
Soybean	1446	1800
Cotton	1272	411
Tobacco	958	892
Wheat	802	895
Yerba Mate	555	474
Sorghum	469	604
Cocoa	264	282
Barley	173	135

Notes: See Appendix A.1 for variable definition and sources. The table presents the number of Brazilian municipalities that produce each of the crop in our analysis.

sures." Intermediate-level inputs/advanced management: "Under an intermediate level of input (improved management assumption), the farming system is partly market oriented. Production for subsistence plus commercial sale is a management objective. Production is based on improved varieties, on manual labor with hand tools and/or animal traction and some mechanization, is medium labor intensive, uses some fertilizer application and chemical pest disease and weed control, adequate fallows and some conservation measures." High-level inputs/advanced management: "Under the high input, advanced management assumption, the farming system is mainly market oriented. Commercial production is a management objective. Production is based on improved high yielding varieties, is fully mechanized with low labor intensity and uses optimum applications of nutrients and chemical pest, disease and weed control."

International commodity prices: Data on international commodity price comes from the Global Economic Monitor (GEM) Commodities database published by the World Bank, which is publicly available on their website.

Outcome Variables

Total GDP per capita: Data on Total Gross Domestic of Municipalities is from the *Produto Interno Bruto dos Municípios* series by IBGE. We use the values for 2000 and 2010 in our analysis. The values are reported in current values in Reais in the respective years. In order to deflate all 2000 values to 2010 Reais, we used the annual IPCA (Índice Nacional de Preços ao Consumidor Amplo). In order to obtain per capita values, we divide the real values by the total population in each municipality taken from the Population Census.

Agricultural GDP per capita: Data on Agricultural GDP per Capita is obtained in the same way as Total GDP per Capita

Manufacturing GDP per capita: Data on Agricultural GDP per Capita is constructed in the same way as Total and Agricultural GDP per Capita.

Services GDP per capita: Data on Agricultural GDP per Capita is constructed in the same way as Total, Agricultural and Manufacturing GDP per Capita.

Employment share in agriculture: Data on employment share in agriculture come from the Brazilian Population Censuses of 2000 and 2010. The variable is calculated as total number of people who reported working in the agricultural sector divided by the total number of people that reported being employed in any sector of the economy. Sector of employment is reported according to the first version of the CNAE-Domiciliar in both 2000 and 2010. The agricultural sector is defined as any sector with the codes ranging from 01 to 05, which encompasses the Sections A: Agriculture, Livestock, Silviculture and Forest Production and B: Fishing. Original data come at the individual level and we use the municipality-level aggregation provided by the Atlas do Desenvolvimento Humano no Brasil (PNUD, 2020).

Employment share in manufacturing: Data on employment share in manufacturing come from the Brazilian Population Censuses of 2000 and 2010. The variable is calculated and aggregated in the same way as employment share in agriculture. The manufacturing sector is defined as Sections C: Extractive Industries and D: Transformation Industries, with the codes ranging from 10 to 37.

Employment share in services: Data on employment share in services come from the Brazilian Population Censuses of 2000 and 2010. The variable is calculated and aggregated in the same way as employment share in agriculture and manufacturing. The service sector is defined as Sections G: Commerce; H: Transportation; I: Accommodation and Food; J: Information and Communication; K: Financial Activities; L: Real Estate; M: Professional, Scientific and Technical Activities; N: Administrative Activities; O: Public administration, defense and social security; P: Education; Q: Health; R: Arts, Culture, Sport and Recreation; S: Other Services; T: Domestic Services, with the codes ranging from 45 to 97.

Wages in agriculture: Data on the wage in agriculture come from the Brazilian Population Censuses of 2000 and 2010. All wages are reported in current values in Reais in the respective years of the two Censuses. In order to deflate all 2000 values to 2010 Reais, we used the annual IPCA (Índice Nacional de Preços ao Consumidor Amplo).

Wages in manufacturing: Data on the wage in manufacturing come from the Brazilian Population Censuses of 2000 and 2010. The variable is constructed in the same way as the wages in agriculture.

Wages in services: Data on the wage in services come from the Brazilian Population Censuses of 2000 and 2010. The variable is constructed in the same way as the wages in agriculture and manufacturing.

Urban population share: Data on the urban population share come from the Brazilian Population Censuses of 2000 and 2010. The variable is constructed by dividing the total population living in urban areas with the total population in each municipality.

Output in agroindustry: Data on the output value in the agroindustry is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA.[29] Data of 2006 is from Tabela 4081 and of 2017 from Tabela 6960.

Employment share in manufacturing - RAIS: Data on alternative employment share in manufacturing come from the Relação Anual de Informações Sociais (RAIS) from the Brazilian Ministry of Labor.[30] The manufacturing sector is defined the same way as before, using Sections C: Extractive Industries and D: Transformation Industries from CNAE, with the codes ranging from 10 to 37.

Employment share in agroindustry: Data on employment share in agroindustry come from the Relação Anual de Informações Sociais (RAIS) from the Brazilian Ministry of Labor. The agroindustry sector is defined using codes 152 to 214, which include the processing of food; beverage production; manufacturing of textile products; clothing; wood and paper processing.

Employment share in heavy industry: Data on employment share in agroindustry come

[29] www.sidra.org.br.

[30] As of 2020, the Ministry of Labor has been incorporated into the Ministry of Economy.

from the Relação Anual de Informações Sociais (RAIS) from the Brazilian Ministry of Labor. The heavy industry sector is defined using codes 24 to 37, which include the chemical industry; pharmaceuticals; metallurgy; machinery; electronics; automotive industry; furniture.

Total farmland: Data on the total farmland is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. Data of 2006 is from Tabela 787 and of 2017 from Tabela 6878.

Machine intensity: Data on the machine intensity is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. We construct the variable by obtaining the total number of tractors, seeders, planters, harvesters and fertilizers used in the agricultural establishments (farms) in each municipality. We then divide the total number of machines by the total farmland. Data of 2006 is from Tabela 860 (tractors) and Tabela 861 (other machines). Data of 2017 is from Tabela 6898 (tractors) and Tabela 6872 (other machines).

Share of land with GE seeds: Data on the share of land harvested with GE seeds is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. Data of 2006 is from Tabela 824 and of 2017 from Tabela 6958.

Farm productivity: Data on farm productivity is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. We construct the variable by dividing the total value of output in agricultural for each municipality by the total farmland. Data of 2006 for Total value of production in agriculture is from Tabelas: 1823 (value of production by seasonal crop), 1177/1178 (value of production by permanent crop), 818 (value of production by horticulture product), 815 (value of production by forestry product), 816 (value of production by vegetable extraction product), 782 (value of bovines), 937 (value of swines), 943 (value of poultry). Data for 2017 is from Tabelas: 6957 (value of production by seasonal crop), 6955 (value of production by permanent crop), 6953 (value of production by horticulture product), 6949 (value of production by vegetable extraction product), 6947 (value of production by forestry product), 6940 (value of poultry), 6927 (value of swines), 6910 (value of bovines).

Output per worker: Data on output per worker is from the Brazilian Agricultural Census of 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. We construct the variable by dividing the total value of output in agricultural for each municipality by the number of workers in agriculture. Data for the number of workers in 2006 is from Tabela 956 and in 2017 from Tabela 6889.

Land Gini: Data on Land Gini is from the Brazilian Agricultural Census of 1995, 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. We detail the construction of the variable in Appendix A.2. We use data for the total area and number of farms in each municipality. Data for the area of agricultural establishments in 1995 is from Tabela 315, of 2006 is from Tabela 787 and of 2017 from Tabela 6878. Data for the number of farms in 1995 is from Tabela 312, of 2006 is from Tabela 860 and of 2017 from Tabela 6880.

Area of farms $\leq$ 1.000 ha: Data on the area of farms with more than 1.000 ha is from the Brazilian Agricultural Census of 1995, 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. The data is from the same tables used in the Land Gini.

Number of farms $\leq$ 1.000 ha: Data on the number of farms with more than 1.000 ha is from the Brazilian Agricultural Census of 1995, 2006 and 2017 and it is sourced from the IBGE online data repository SIDRA. The data is from the same tables used in the Land Gini.

Control Variables

Share of rural population: Data on the share of rural population come from Brazilian Population Censuses of 1991. The variable is constructed by dividing the total number of people that reported living in rural area with the total population in each municipality. Original data come at the individual level and we use the municipality-level aggregation provided by the Atlas do Desenvolvimento Humano no Brasil (PNUD, 2020).

Illiteracy rate: Data on the illiteracy come from Brazilian Population Censuses of 1991. The variable is calculated considering only people 10 years or older, as the total number of people who is able to read and write divided by the total number of people. Original data come at the individual level and we use the municipality-level aggregation provided by the Atlas do Desenvolvimento Humano no Brasil (PNUD, 2020).

Income per capita: Data on income per capita come from Brazilian Population Censuses of 1991. Income is defined for every person as the sum of income coming from all sources. Original data come at the individual level and we use the municipality-level aggregation provided by the Atlas do Desenvolvimento Humano no Brasil (PNUD, 2020).

Population density: Data on population density come from Brazilian Population Censuses of 1991 and IpeaData. We construct the variable by dividing the total population of each municipality by its total area in square kilometers. Data of municipality area is from

IpeaData.

Variables Used in Figures and Additional Results

Employment shares from Figure 3: Data on employment shares since 1960 in Brazil come from IPUMS International microdata for the Brazilian Population Census (Min, 2019). The dataset enables us to construct employment shares in the usual way, but also keeping the variables in each year harmonized throughout the different definitions of occupations in the Censuses using ISCO and SIC codes, which follow a similar structure than that of CNAE.

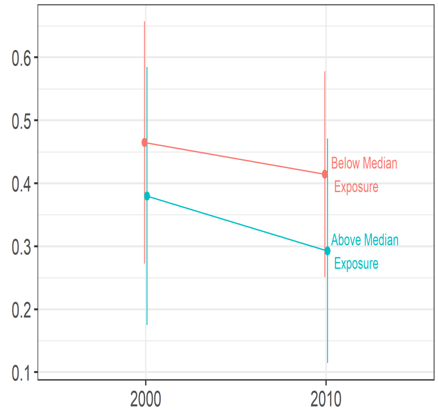
A.2 Land Gini Calculation

The three rounds of the Agricultural Census (1995, 2006 and 2017) divide the area of rural properties into different intervals, reporting the total number and total area of properties across different bins of total area. We are then able to calculate the mean property size within each interval. We follow Nunn (2008) and use the Stata program *ineqdec0* written by Stephen P. Jenkins, to obtain the Gini coefficient as follows:

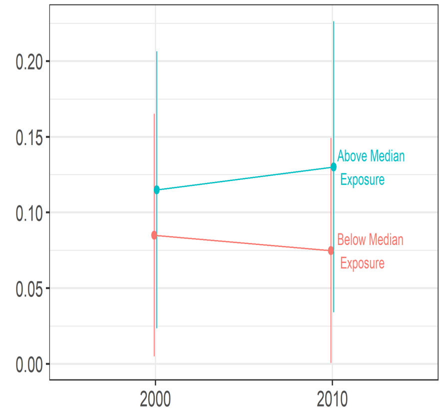
$$\text{Gini} = 1 + \left(\frac{1}{n}\right) - \frac{2\sum_{1}^{n}(n-i+1)\,a_i}{n\sum_{i}^{n}a_i}$$

where n is the number of farms, a_i is farm size in acres, and i denotes the rank in ascending order of a_i.

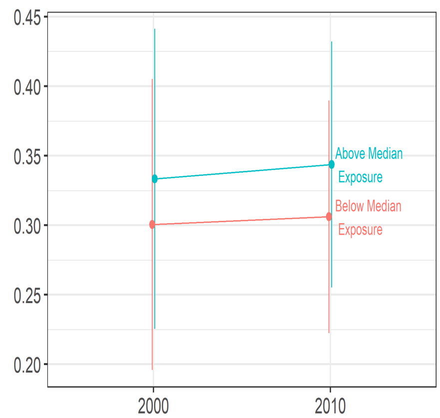
B Additional Figures



(a) Agriculture



(b) Manufacturing



(c) Services

Figure B1: Employment Shares (Mean and 1SD Error Bars)

Notes: The figures display the average employment shares, and their respective 1 standard error bars, in Brazilian municipalities that were ranked above and below the ΔCE measure. Data are from the 2000 and 2010 Population Census. See Appendix A.1 for variable definition and sources.

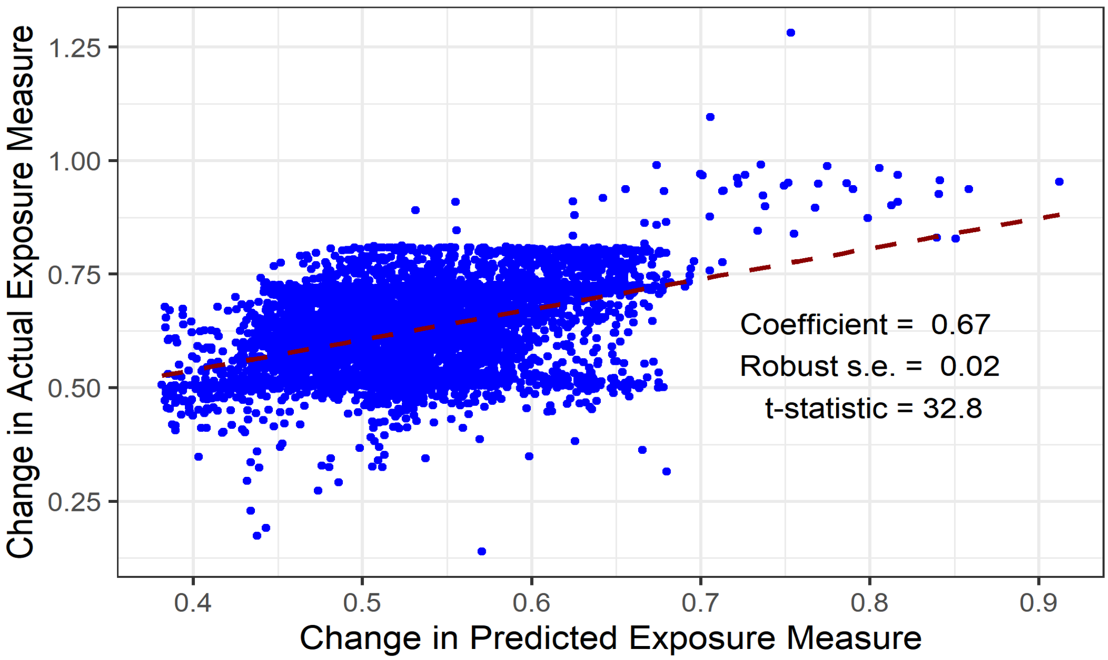


Figure B2: Actual and Predicted Exposure Measure

Notes: The figure displays the scatter plot of actual and predicted ΔCE measure constructed in Equation 3. Predicted shares are obtained from the FML model. The figure reports the coefficient, robust standard error, and t-statistic of the bivariate regression between the two variables. See Appendix A.1 for variable definition and sources.

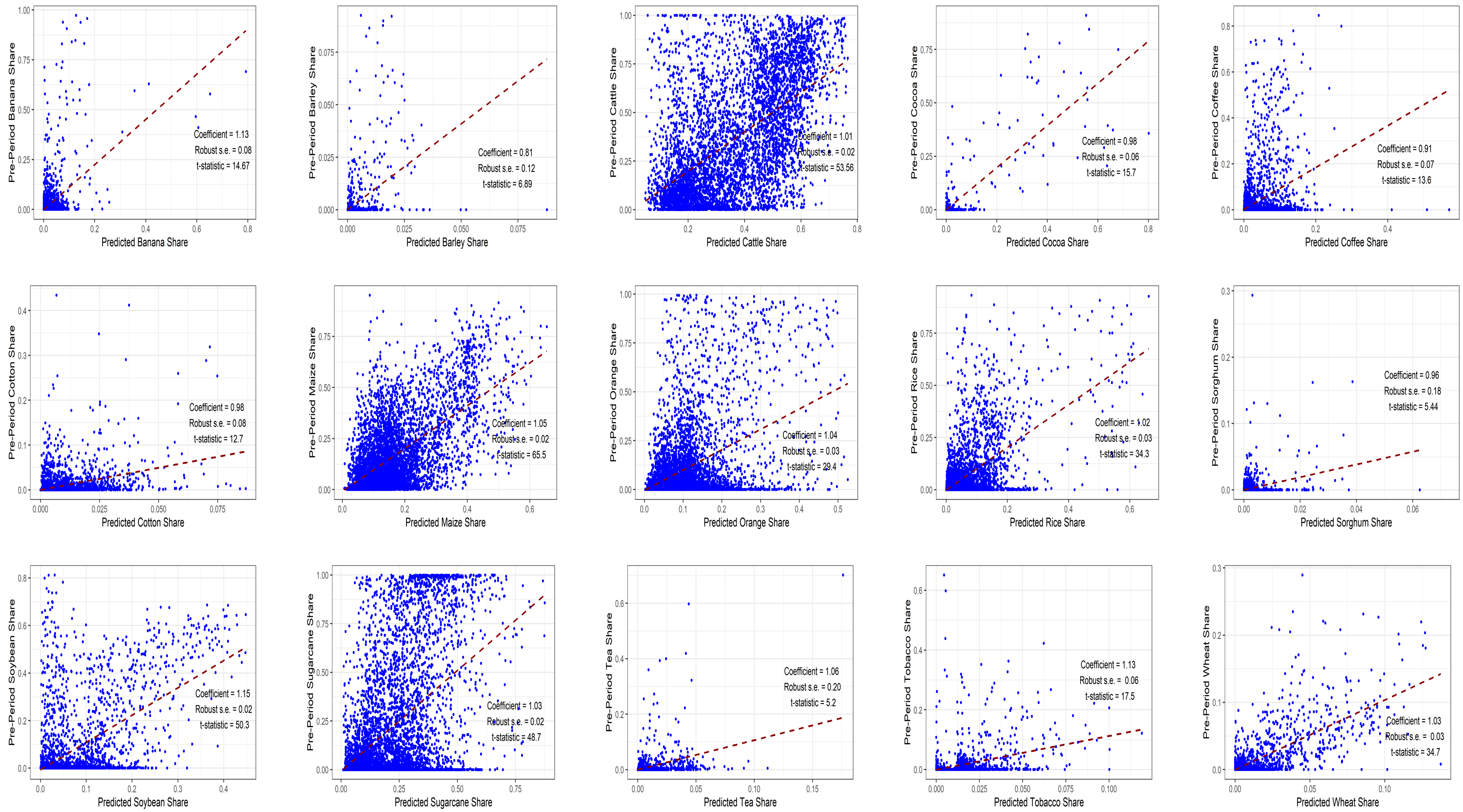


Figure B3: Actual and Predicted Agricultural Shares

Notes: The figures display the scatter plot of actual and predicted shares for all crops in the sample. The figure reports the coefficients, robust standard errors, and t-statistics of the bivariate regression between the two variables in each graph. Predicted shares are obtained from the FML model. See Appendix A.1 for variable definition and sources.

C Additional Results

Table C1: The Effect of the Commodity Shock on Employment Shares in the Formal Sector

	(1)	(2)	(3)	(4)
Panel A.	Δ Employment Share in Agriculture			
Δ CE	−0.355***	−0.353***	−0.176***	−0.195***
	(0.033)	(0.040)	(0.052)	(0.056)
Adj. R^2	0.026	0.026	0.050	0.171
Panel B.	Δ Employment Share in Manufacturing			
Δ CE	0.135***	0.067**	0.051	0.098**
	(0.029)	(0.032)	(0.039)	(0.045)
Adj. R^2	0.009	0.012	0.044	0.059
Panel C.	Δ Employment Share in Services			
Δ CE	0.222***	0.289***	0.155**	0.128*
	(0.043)	(0.055)	(0.068)	(0.074)
Adj. R^2	0.012	0.013	0.052	0.062
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,455	5,455	5,455	5,455

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2000-2010 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, and C are -0.023 (0.125), 0.001 (0.117), and 0.021 (0.170). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table C2: The Effect of the Commodity Shock on Total Employment Across Sectors

	(1)	(2)	(3)	(4)
Panel A.	Δ Log Total Employment			
Δ CE	0.442***	0.155***	0.198***	0.107**
	(0.045)	(0.048)	(0.071)	(0.053)
Adj. R^2	0.016	0.060	0.157	0.124
Panel B.	Δ Log Total Employment in Agriculture			
Δ CE	−0.076	−0.267***	−0.215***	−0.279***
	(0.073)	(0.076)	(0.083)	(0.084)
Adj. R^2	0.000	0.008	0.033	0.039
Panel C.	Δ Log Total Employment in Manufacturing			
Δ CE	2.238***	2.075***	1.562***	1.449***
	(0.144)	(0.156)	(0.167)	(0.172)
Adj. R^2	0.046	0.047	0.061	0.064
Panel D.	Δ Log Total Employment in Services			
Δ CE	0.523***	0.156	0.326**	0.392**
	(0.089)	(0.106)	(0.129)	(0.157)
Adj. R^2	0.014	0.020	0.050	0.060
Rural Share in 1991		✓	✓	✓
Region FE			✓	✓
Baseline Controls				✓
Observations	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents estimates of Equation 6. The dependent variable is the 2000-2010 change in the listed outcomes. Column (1) reports the estimates without any control. Columns (2)-(4) extend the set of controls by including the share of population living in the rural area; region fixed effect for the 5 macroregions of Brazil; log income per capita; log population density; and illiteracy rate. All municipalities controls are from the population census of 1991. The means and standard deviations (in parentheses) of the dependent variables in panels A, B, C and D are 0.190 (0.220), -0.009 (0.342), 0.173 (0.666), and 0.927 (0.818). Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table C3: The Effect of the Commodity Shock on Structural Transformation - Main Specification Coefficients

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Δ CE	−0.120***	0.151***	0.001	0.032**	0.444***	−0.300**	0.581***
	(0.018)	(0.014)	(0.013)	(0.016)	(0.093)	(0.136)	(0.096)
Log Income per Capita in 1991	0.018***	−0.005*	−0.009***	0.006*	−0.010***	−0.056**	0.013
	(0.004)	(0.003)	(0.003)	(0.003)	(0.020)	(0.029)	(0.021)
Log Population Density in 1991	0.0004	−0.001*	−0.001	0.0007	−0.021***	−0.006	−0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.007)	(0.006)
Share of Rural Population in 1991	−0.054***	−0.004	0.062***	0.094***	−0.196***	−0.114***	0.006
	(0.006)	(0.004)	(0.004)	(0.006)	(0.030)	(0.042)	(0.030)
Illiteracy Rate in 1991	−0.0007***	0.0002*	0.0003***	0.00008	−0.0008	−0.002	−0.002**
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.001)	(0.001)	(0.001)
Northeast Region Dummy	0.001	0.016***	−0.012***	0.00003	0.136***	0.010**	−0.024
	(0.006)	(0.004)	(0.004)	(0.004)	(0.028)	(0.041)	(0.030)
Southeast Region Dummy	−0.010*	0.029***	−0.004	0.009**	0.163***	0.028	0.065**
	(0.006)	(0.004)	(0.003)	(0.004)	(0.024)	(0.036)	(0.027)
South Region Dummy	−0.019***	0.032***	0.003	0.013***	0.182***	0.026	0.182***
	(0.006)	(0.004)	(0.004)	(0.005)	(0.027)	(0.038)	(0.029)
Central-West Region Dummy	−0.0112*	0.025***	−0.011***	−0.006	0.195***	0.006	0.379***
	(0.006)	(0.004)	(0.004)	(0.005)	(0.028)	(0.040)	(0.029)
Adj. R^2	0.086	0.052	0.176	0.127	0.025	0.006	0.145
Observations	5,475	5,475	5,475	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Tables 4 and 5, reporting all the coefficients for the controls. The North region dummy is omitted in the set of region dummies. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C4: The Effect of the Commodity Shock on Structural Transformation - Additional Controls

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Δ CE	−0.101***	0.125***	0.006	0.0294	0.293***	−0.214	0.434***
	(0.020)	(0.016)	(0.014)	(0.018)	(0.108)	(0.158)	(0.112)
Adj. R^2	0.113	0.107	0.167	0.154	0.057	0.015	0.209
Observations	4,865	4,865	4,865	4,865	4,865	4,865	4,865

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 with additional controls. We now add both the baseline controls reported in column (4) of Tables 4 and 5 and the following variables: share of women in the labor force in 1991; log value of output per total farmland in 1995; altitude; state capital dummy; distance to the state capital; distance to the federal capital; dummies for individual and groups of commodities. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C5: The Effect of the Commodity Shock on Input Use and Land Inequality - Additional Controls

	Δ Log Total Farmland	Δ Log Machine Intensity	Δ Share of Land with GE Seeds	Δ Land Gini (2006-2017)	Δ Land Gini (1995-2017)
Δ CE	0.030	0.551**	0.133***	0.092***	0.116***
	(0.125)	(0.271)	(0.034)	(0.024)	(0.029)
Adj. R^2	0.078	0.035	0.260	0.089	0.137
Observations	4,841	4,841	4,841	4,841	4,841

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 with additional controls. We now add both the baseline controls reported in column (4) of Tables 7 and 9 and the following variables: share of women in the labor force in 1991; log value of output per total farmland in 1995; altitude; state capital dummy; distance to the state capital; distance to the federal capital; dummies for individual and groups of commodities. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C6: The Effect of the Commodity Shock on Structural Transformation - Alternative Shares for the Exposure Measure

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Δ CE - Actual Shares	−0.159***	0.207***	−0.008	0.042***	0.526***	−0.248*	0.814***
	(0.019)	(0.014)	(0.013)	(0.016)	(0.098)	(0.138)	(0.096)
Adj. R^2	0.091	0.070	0.176	0.127	0.027	0.006	0.150
Δ CE - No Cattle	−0.159***	0.167***	0.004	0.019	0.715***	−0.367**	0.964***
	(0.023)	(0.017)	(0.016)	(0.021)	(0.107)	(0.167)	(0.116)
Adj. R^2	0.088	0.048	0.176	0.126	0.028	0.006	0.149
Δ CE - High Inputs	−0.111***	0.149***	−0.004	0.021	0.346***	−0.264**	0.409***
	(0.018)	(0.013)	(0.012)	(0.016)	(0.087)	(0.130)	(0.091)
Adj. R^2	0.086	0.053	0.176	0.127	0.024	0.006	0.142
Observations	5,475	5,475	5,475	5,475	5,475	5,475	5,475

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Tables 4 and 5 substituting the shares used in the construction of the Δ CE measure with: the mean production of each crop during the ten years between 2000 and 2010 (Actual Shares); the pre-period shares excluding cattle (No Cattle); the high inputs potential yields from FAO-GAEZ. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C7: The Effect of the Commodity Shock on the Land Use and Capital Inputs - Alternative Shares for the Exposure Measure

	Δ Log Total Farmland	Δ Log Machine Intensity	Δ Share of Land with GE Seeds
Δ CE - Actual Shares	−0.185*	0.212	0.279***
	(0.108)	(0.213)	(0.036)
Adj. R^2	0.063	0.021	0.198
Δ CE - No Cattle	0.074	1.687***	0.322***
	(0.134)	(0.316)	(0.042)
Adj. R^2	0.062	0.029	0.197
Δ CE - High Inputs	−0.014	0.670***	0.135***
	(0.104)	(0.220)	(0.033)
Adj. R^2	0.062	0.023	0.192
Observations	5,446	5,446	5,446

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 7 substituting the shares used in the construction of the Δ CE measure with: the mean production of each crop during the ten years between 2000 and 2010 (Actual Shares); the pre-period shares excluding cattle (No Cattle); the high inputs potential yields from FAO-GAEZ. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C8: The Effect of the Commodity Shock on Land Inequality - Alternative Shares for the Exposure Measure

	Δ Land Gini (2006-2017)	Δ Land Gini (1995-2017)
Δ CE - Actual Shares	0.082***	0.123***
	(0.020)	(0.025)
Adj. R^2	0.075	0.111
Δ CE - No Cattle	0.067***	0.077**
	(0.025)	(0.032)
Adj. R^2	0.073	0.108
Δ CE - High Inputs	0.053***	0.075***
	(0.019)	(0.024)
Adj. R^2	0.073	0.109
Observations	5,446	4,919

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 9 substituting the shares used in the construction of the Δ CE measure with: the mean production of each crop during the ten years between 2000 and 2010 (Actual Shares); the pre-period shares excluding cattle (No Cattle); the high inputs potential yields from FAO-GAEZ. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C9: The Effect of the Commodity Shock on Employment Shares - Observations at the Micro-Region Level

	Δ Employment Share Agriculture	Δ Employment Share Manufacturing	Δ Employment Share Services	Δ Urban Population Share	Δ Log Wages Agriculture	Δ Log Wages Manufacturing	Δ Log Wages Services
Δ CE	-0.169***	0.143***	0.046*	0.046	0.432**	-0.264	0.246
	(0.040)	(0.029)	(0.024)	(0.035)	(0.173)	(0.199)	(0.199)
Adj. R^2	0.170	0.105	0.392	0.242	0.131	0.071	0.345
Observations	558	558	558	558	558	558	558

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Tables 4 and 5 and aggregates the municipality-level variables at the microregion-level. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C10: The Effect of the Commodity Shock on the Land Use and Capital Inputs - Observations at the Microregion Level

	Δ Log Total Farmland	Δ Log Machine Intensity	Δ Share of Land with GE Seeds
Δ CE	-0.223	1.137***	0.201***
	(0.246)	(0.428)	(0.078)
Adj. R^2	0.129	0.105	0.265
Observations	557	557	557

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 7 and aggregates the municipality-level variables at the microregion-level. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.

Table C11: The Effect of the Commodity Shock on Land Inequality - Observations at the Micro-Region Level

	Δ Land Gini (2006-2017)	Δ Land Gini (1995-2017)
Δ CE	0.028	0.116*
	(0.048)	(0.069)
Adj. R^2	0.194	0.262
Observations	557	557

Notes: See Appendix A.1 for variable definition and sources. The table presents then main specification estimates of Equation 6 as reported in column (4) of Table 9 and aggregates the municipality-level variables at the microregion-level. Robust standard errors reported in parentheses. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10%.