Technical Appendix to Agricultural Productivity and Structural Transformation: Not for Publication

Paula Bustos, Bruno Caprettini and Jacopo Ponticelli

This technical appendix contains a more detailed exposition of the theoretical framework presented in the paper, a detailed description of the variables used in the empirical analysis as well as additional supplementary material for the paper. The appendix follows the same structure of the paper. Section I provides additional material on section I in the paper. In particular, it provides a detailed description of the data sources used in the Figures presented in section I. Section II contains the proofs of the propositions stated in section II of the paper, and an extension of the model presented in the paper that includes a non-tradable sector. Section III provides a detailed description of all the variables used in the empirical analysis. In section IV we report additional empirical results discussed in section IV of the paper. Finally, in section V we report the tables of a set of robustness checks described in the paper.

I. Agriculture in Brazil

In this section we provide additional information on the data used to produce Figure 1 and Table 1 in the paper, which show stylized facts about agriculture in Brazil.

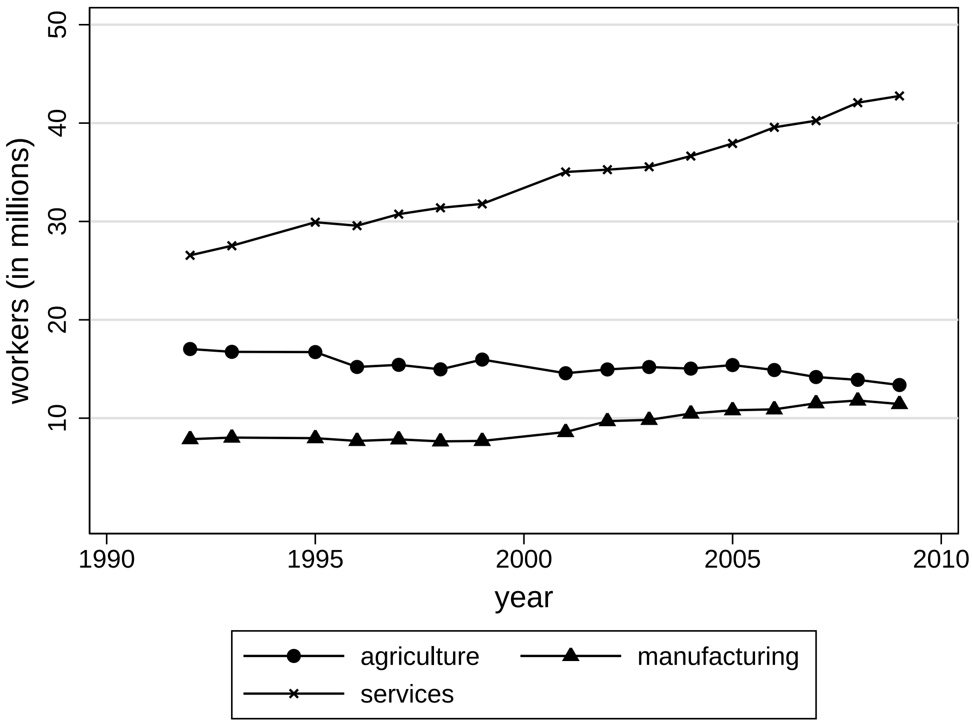
Figure 1. The data used on soy output, area planted with soy and area planted with maize in first and second season is from CONAB, while data on number of workers employed in soy production is from PNAD. CONAB, Companhia Nacional de Abastecimento, is an agency within the Brazilian Ministry of Agriculture. CONAB carries out monthly surveys to monitor the evolution of the harvest of all major crops in Brazil: the surveys are representative at state level and are constructed by interviewing on the ground farmers, agronomists and financial agents in the main cities of the country. All data can be downloaded at: `http://www.conab.gov.br/conteudos.php?a=1252&t=`. PNAD is a national household survey representative at state level carried out yearly by the IBGE (the survey was not carried out in 1994 and in the census years: 1991, 2000 and 2010). Since the PNAD coverage changed over time, to harmonize the sample across years we exclude from the PNAD data: (i) workers located in the states of: Rondonia, Acre, Amazonas, Roraima, Pará and Amapá (North macro-region) because only urban areas (and not rural areas) of these states were covered until 2004; (ii) workers located in the states of: Tocantins, Mato Grosso do Sul, Goias and the Distrito Federal because the sample of households in these states is not complete in the years from 1992 to 1997. We harmonized data from CONAB with the PNAD coverage such that numerator and denominator are constructed using the same subset of states.

Notice that the decrease in labor intensity in soy production between 1996 and 2006 implied by Figure 1(c) is larger than the one reported in the text and Table 1. This is because the data sources are different. Figure 1(c) displays yearly data on area planted with soy from the CONAB survey and yearly data on employment in soy production from the PNAD survey. Table 2 instead is based on data on area planted and employment from the Agricultural Censuses of 1996 and 2006. This generates two sources of discrepancy in the reported labor intensity data. First, Figure 1(c) excludes some states due to limited coverage of the PNAD survey. Second, the unit of observation in PNAD is the individual while the Agricultural Census surveys farms. As a result, labor intensity in soy production in Table 1 is computed as total land in farms whose main activity is soy divided by total number of workers in farms whose main activity is soy, which tends to overestimate the number of workers in soy whenever farms whose main activity is soy produce also other crops which are, on average, more labor intensive.

Table 1. The data on land use and labor intensity in agriculture is from the Brazilian Agricultural Censuses of 1996 and 2006. The data has been downloaded from the IBGE Sidra repository (last update: November 2015). In particular: data on land use by principal activity (total, permanent crops, seasonal crops, cattle ranching, forestry and non-usable land) is from Table 317 for 1996 and Table 1011 for 2006. Data on land use for specific seasonal crops (soy, cereals and other seasonal crops) is from Table 501 for 1996 and from Table 822 for 2006. Cereals is the sum of: rice, maize and wheat. To compute labor intensity we use data on land in farms by principal activity of the farm and number of workers in farms by principal activity of the farm. Data on land in farm by principal activity is from Table 491 for 1996 and from Table 787 in 2006. Data on land in farm by specific seasonal crops (soy, cereals and other seasonal crops) is from Table 1 for 1996 (available at `ftp://ftp.ibge.gov.br/Censos/Censo_Agropecuario_1995_96/Brasil/`) and from Table 1013 for 2006. Data on total number of workers in farms by principal activity of the farm (permanent crops, seasonal crops, cattle ranching and forestry) is from Table 321 for 1996 and from Table 956 for 2006. Data on number of workers for specific seasonal crops (soy, cereals and other seasonal crops) is from Table 1013 for 1996 and from Table 806 and 916 for 2006. Notice that the definition of "principal activity" of the farm changed between the Agricultural Census of 1996 and the one of 2006. In 1996 higher specialization was required for farms to be classified under one of the categories reported, and those that did not produce at least 2/3 of the value within a single category were classified under the "mixed activity" category. In 2006 farms were classified according to the activity that accounted for the simple majority of production and no "mixed activity" category existed.

Figure A1 shows the evolution between 1992 and 2009 of the total number of workers (expressed in million) employed in different sectors of the Brazilian economy.

Figure A1 Employment in agriculture, industry, services (1992-2009)



Notes: Services include: commerce, lodging and restaurants, transport, finance, housing services, other personal services, domestic workers and construction. Data from PNAD.

II. Model

In this section, we provide derivations of all expressions and results reported in the paper.

A. *Setup*

In this section we show that labor-augmenting technical change reduces the marginal product of labor when the elasticity of substitution is smaller than the land share of output.

First, we obtain $MPL_a = \frac{\partial Q_a}{\partial L_a}$ by differentiating the agricultural production function described by equation (1) w.r.t. labor:

$$MPL_a = A_N \Theta^{\frac{\sigma}{\sigma-1}-1} \gamma L_a^{\frac{\sigma-1}{\sigma}-1} A_L^{\frac{\sigma-1}{\sigma}} \tag{A.1}$$

where, to save space, we define $\Theta \equiv \gamma \left(A_L L_a\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(A_T T\right)^{\frac{\sigma-1}{\sigma}}$. Next, we obtain $\frac{\partial MPL_a}{\partial A_L}$ as follows:

$$\frac{\partial MPL_a}{\partial A_L} = A_N \Theta^{\frac{1}{\sigma-1}} \gamma L_a^{\frac{-1}{\sigma}} A_L^{\frac{-1}{\sigma}} \frac{\sigma-1}{\sigma} \left[1 + \frac{1}{\sigma-1} \Theta^{-1} \gamma \left(A_L L_a\right)^{\frac{\sigma-1}{\sigma}}\right].$$

Then, when $\sigma < 1$, $\frac{\sigma-1}{\sigma} < 0$. As a result, $\frac{\partial MPL_a}{\partial A_L} < 0$ iff the last term in brackets is positive, which is true as long as σ satisfies the following condition:

$$\sigma < \frac{(1-\gamma)\left(A_T T\right)^{\frac{\sigma-1}{\sigma}}}{\gamma\left(A_L L_a\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(A_T T\right)^{\frac{\sigma-1}{\sigma}}} = \frac{T_a MPT_a}{Q_a} \equiv 1 - \Gamma. \tag{A.2}$$

Note that this condition is a function of the equilibrium level of employment in agriculture. In particular, in the relevant case where $\sigma < 1$ the land share $(1-\Gamma)$ is increasing on the level of agricultural employment. As a result, the condition is more likely to be satisfied when the equilibrium level of agricultural employment is high.

B. *Equilibrium*

In this section we detail the derivations necessary to obtain the equilibrium level of employment in agriculture. First, note that profit maximization implies that the value of the marginal product of labor must equal the wage in both sectors: $P_a MPL_a = w = P_m MPL_m$. Thus, the marginal product of labor in agriculture is determined by international prices and manufacturing productivity:

$$MPL_a = \left(\frac{P_m}{P_a}\right)^* A_m. \tag{A.3}$$

Next, starting from equation (A.1), and using the land market clearing condition ($T_a = T$), MPL_a can be written as follows:

$$MPL_a = A_N A_L \gamma \left[\gamma + (1-\gamma)\left(\frac{A_T T}{A_L L_a}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{1}{\sigma-1}}. \tag{A.4}$$

Then, (A.3) and (A.4) can be used to solve for the equilibrium level of employment in agriculture:

$$L_a^* = \frac{A_T T}{A_L}\left\{\frac{\gamma}{1-\gamma}\left[\frac{1}{\gamma^\sigma}\left(\frac{P_a}{P_m}\frac{A_N A_L}{A_m}\right)^{1-\sigma} - 1\right]\right\}^{\frac{\sigma}{1-\sigma}}. \tag{A.5}$$

Next, we obtain equation (4), where L_a^* is written as a function of the equilibrium land share. We calculate the ratio of the labor to land share using the production function described by equation (1):

$$\frac{\Gamma}{1-\Gamma} \equiv \frac{\frac{L_a MPL_a}{Q_a}}{\frac{T_a MPT_a}{Q_a}} = \frac{\gamma}{(1-\gamma)}\left(\frac{A_L L_a}{A_T T}\right)^{\frac{\sigma-1}{\sigma}}. \tag{A.6}$$

Note that the equalization of the value of the marginal product of labor across sectors, described by equations (A.3) and (A.4), implies:

$$\frac{(1-\gamma)}{\gamma}\left(\frac{A_T T}{A_L L_a}\right)^{\frac{\sigma-1}{\sigma}} = \frac{1}{\gamma}\left(\frac{P_m}{P_a}\frac{A_m}{A_N \gamma A_L}\right)^{\sigma-1} - 1. \tag{A.7}$$

Equations (A.6) and (A.7) can be used to obtain a solution for the the equilibrium land share. This solution implies that the parameter restriction necessary for labor augmenting technical change to be strongly labor saving is:

$$\sigma < 1 - \Gamma^* = 1 - \gamma\left(\frac{P_m}{P_a}\frac{A_m}{A_N \gamma A_L}\right)^{1-\sigma}. \tag{A.8}$$

Note that as σ gets closer to one, $1-\Gamma^*$ converges to the parameter $1-\gamma$. In the relevant case where $\sigma < 1$, the land share is increasing in world relative prices for agricultural goods and relative agricultural productivity. Thus, this condition is more likely to be satisfied in regions with a comparative advantage in agriculture.

C. *Technological Change and Structural Transformation*

In this section we analyze the response of agricultural and manufacturing employment to labor-augmenting technical change.

Labor-augmenting technical change

The effect of labor augmenting technical change on agricultural employment depends on whether the elasticity of substitution is smaller than the equilibrium land share of agricultural production ($\sigma < 1-\Gamma^*$). When this condition is satisfied, we say that land and labor are strong complements.

a) Land and labor are strong complements: $\frac{\partial L_a^*}{\partial A_L} < 0$ and $\frac{\partial L_m^*}{\partial A_L} > 0$.

Proof. We obtain $\frac{\partial L_a^*}{\partial A_L}$ from the solution for L_a^* as a function of the equilibrium land share given by equation (4), as follows:

$$\begin{aligned}\frac{\partial L_a^*}{\partial A_L} &= \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{\sigma}{1-\sigma}} \frac{A_T T}{A_L^2} \left\{ \frac{\sigma}{1-\sigma} \frac{\partial \frac{1-\Gamma^*}{\Gamma^*}}{\partial A_L} \frac{A_L}{\frac{1-\Gamma^*}{\Gamma^*}} - 1 \right\} \\ &= \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{\sigma}{1-\sigma}} \frac{A_T T}{A_L^2} \left\{ \frac{\sigma}{1-\Gamma^*} - 1 \right\}\end{aligned} \tag{A.9}$$

where the last step used the following solution for the elasticity of the land to labor share ratio with respect to labor-augmenting technical change:

$$\frac{\partial \frac{1-\Gamma^*}{\Gamma^*}}{\partial A_L} \frac{A_L}{\frac{1-\Gamma^*}{\Gamma^*}} = \frac{1-\sigma}{1-\Gamma^*}. \tag{A.10}$$

Equation (A.9) implies that $\frac{\partial L_a^*}{\partial A_L} < 0$ iff $\frac{\sigma}{1-\Gamma^*} < 1$. In this case, $\frac{\partial L_m^*}{\partial A_L} > 0$ because $L_a^* + L_m^* = L$.

b) Land and labor are not strong complements: $\frac{\partial L_a^*}{\partial A_L} > 0$ and $\frac{\partial L_m^*}{\partial A_L} < 0$.

Proof. (A.9) implies $\frac{\partial L_a^*}{\partial A_L} > 0$ iff $\frac{\sigma}{1-\Gamma^*} > 1$. In this case, $\frac{\partial L_m^*}{\partial A_L} < 0$ because $L_a^* + L_m^* = L$.

D. *Empirical Predictions*

Prediction 1. If land and labor are strong complements in production, labor augmenting technical change in agriculture (A_L) :

(a) increases the value of output per worker: $\frac{\partial \frac{P_a^* Q_a^*}{L_a^*}}{\partial A_L} > 0$;

(b) reduces the labor intensity of production: $\frac{\partial \frac{L_a^*}{T}}{\partial A_L} < 0$;

(c) reduces the employment share of agriculture: $\frac{\partial \frac{L_a^*}{L}}{\partial A_L} < 0$;

(d) increases the employment share of manufacturing: $\frac{\partial \frac{L_m^*}{L}}{\partial A_L} > 0$.

Proof

(a) We start by obtaining the average product of labor $\frac{Q_a}{L_a}$ by dividing the agricultural production function described by equation (1) by L_a :

$$\frac{Q_a}{L_a} = A_N A_L \left[\gamma + (1-\gamma) \left(\frac{A_T T_a}{A_L L_a} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

Next, we use equation (A.4) to write $\frac{Q_a}{L_a}$ as a function of the marginal product of labor, as follows:

$$\frac{Q_a}{L_a} = \gamma^{-\sigma} \left(A_N A_L\right)^{1-\sigma} \left(MPL_a\right)^{\sigma}. \tag{A.11}$$

The equation above implies that an increase in A_L must increase the equilibrium average product of labor as long as $\sigma < 1$. This is because the marginal product of labor must remain constant in equilibrium (see condition (A.3) above). Finally, because prices are set in international markets, the value of the average product of labor also increases.

*(b) and (c)*Note that $\frac{\partial L_a^*}{\partial A_L} < 0$, as shown in section C. As a result, $\frac{\partial \frac{L_a^*}{T}}{\partial A_L} < 0$ because the land endowment is fixed; $\frac{\partial \frac{L_a^*}{L}}{\partial A_L} < 0$ and because the labor endowment is fixed.

(d) Note that $\frac{\partial L_m^*}{\partial A_L} > 0$, as shown in section C. As a result, $\frac{\partial \frac{L_m^*}{L}}{\partial A_L} < 0$ because the labor endowment is fixed.

Prediction 2. Land augmenting technical change in agriculture (A_T) :

(a) does not change the value of output per worker: $\frac{\partial \frac{P_a^* Q_a^*}{L_a^*}}{\partial A_T} = 0$;

(b) increases the labor intensity of production: $\frac{\partial \frac{L_a^*}{T}}{\partial A_T} > 0$;

(c) increases the employment share of agriculture: $\frac{\partial \frac{L_a^*}{L}}{\partial A_T} > 0$;

(d) reduces the employment share of manufacturing: $\frac{\partial \frac{L_m^*}{L}}{\partial A_T} < 0$.

Proof

(a) This result is implied by equations (A.3) and (A.11): land augmenting technical change must leave the equilibrium marginal product of labor unchanged, as a result, the average product of labor is also unchanged. Finally, because prices are set in international markets, the value of the average product of labor is also unchanged.

(b), (c) and (d) These results can be shown with arguments similar to the ones used to prove prediction 1 above.

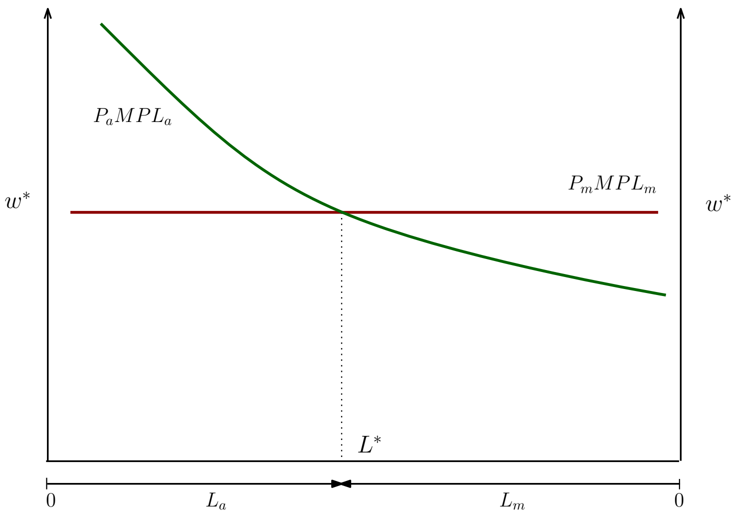
In Figure A2 we show graphically these predictions. Panel (a) of the figure shows the labor market equilibrium. Labor employment is measured on the horizontal axis: employment in agriculture is shown from left to right, while employment in manufacturing from right to left. The length of the axis is determined by the total labor supply, which is fixed. On the vertical axes it is measured the wage and the marginal product of labor in the two sectors. Marginal productivity of labor in manufacturing is represented by the horizontal line in red. Given the agricultural production function (1), marginal productivity of labor is decreasing in agricultural employment. This is represented by a downward sloping curve in the figure (in green). The equilibrium in the labor market (w^*, L^*) is determined by the intersection of the two curves.

Panel (b) of Figure A2 shows graphically predictions (1c) and (1d). When condition (A.2) is satisfied, an increase in the labor augmenting technical parameter A_L shifts downward the marginal product of labor in agriculture. The new curve is shown as the solid green line, which lies below the initial curve (dashed in the figure). In the new equilibrium, labor employment in agriculture falls to maintain the equilibrium on the labor market. Since the labor endowment is fixed, employment share in agriculture also falls, and the employment share in manufacturing grows.

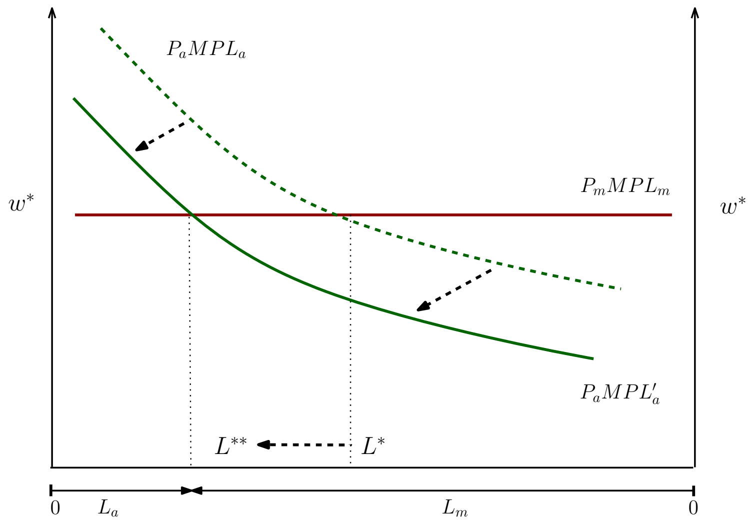
Panel (c) of Figure A2 shows the effect of land augmenting technical change on labor shares. An increase in the land augmenting technical parameter A_T shifts upwards the marginal product of labor in agriculture. The resulting new curve is shown as the solid green line, which lies above the initial curve (dashed in the figure). In the new equilibrium, agricultural labor employment and agricultural labor share fall, while labor employment and labor share in manufacturing grow.

Figure A2
Predictions of the model

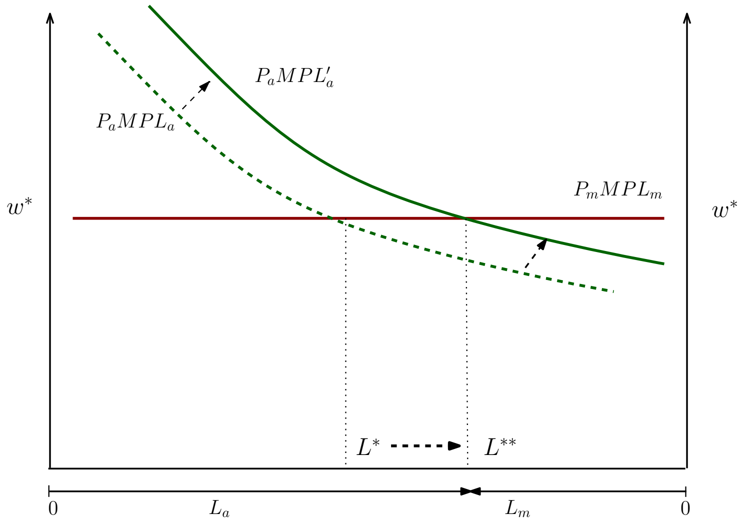
(a) Labor market equilibrium



(b) Labor saving technical change ($\sigma < 1 - \Gamma^*$)



(c) Land augmenting technical change



E. *Services*

In this section we provide an analysis of the model with services.

Setup

The economy is identical to the one described in section A, except that there are three sectors (goods): agriculture (a), manufacturing (m) and services (s). Agriculture and manufacturing are tradable, while services are not.

Production

The production function in services uses only labor and displays constant returns to scale:

$$Q_s = A_s L_s$$

Production functions in agriculture and manufacturing are described by equations 1 and 2, respectively.

Consumption

Because services are non-tradable, production can no longer be determined independently of consumption. Thus, we specify preferences and factor ownership. Consumers have the following Cobb-Douglas preferences over the three goods:

$$U(c_a, c_m, c_s) = c_a^{\alpha_a} c_m^{\alpha_m} c_s^{\alpha_s}. \tag{A.12}$$

where $\alpha_a + \alpha_m + \alpha_s = 1$.[1] There are-two types of agents in the economy: L workers, each endowed with one unit of labor; and T land-owners, each endowed with one unit of land. We assume that workers reside in the same region where they work. In contrast, land owners can reside in any region. We denote by θ the share of land owners residing in the same region where their land is located. Then, aggregate service consumption in a region is $C_s = c_{s,L}\, L + c_{s,T}\, \theta T$, where $c_{s,L}$ is the consumption of workers and $c_{s,T}$ the consumption of land-owners.[2,3]

Equilibrium

The following equilibrium conditions determine the sectorial allocation of employment:

[1] Our use of an homothetic utility function follows the findings in Herrendorf, Rogerson and Valentinyi (2013). They show that an homothetic utility function where the elasticity of substitution across sectors is smaller than one provides the best fit to the Postwar U.S. data when sectoral consumption data is measured in terms of value added. Because we use data on employment to measure structural transformation, our analysis tracks value added better than final goods consumption. As a result we use a non-homothetic utility function. However, we assume that the elasticity of substitution across sectors is equal to one to make the model simpler. We discuss below how the predictions of our model would be modified if this elasticity was smaller than one.

[2] Note that θ is the share of services consumption of land owners that is spent locally. Thus, an alternative interpretation is that land-owners reside locally but consume some services in other regions.

[3] Note that we are not taking into account the local consumption of land owners who reside in the region under consideration but own land in other regions. We consider the role of these absentee landowners at the end of this theoretical section.

1. Profit maximization implies that if there is positive employment in all sectors in equilibrium, the value of the marginal product of labor is equalized across the three sectors:

$$P_a MPL_a = w = P_m MPL_m = P_s MPL_s. \tag{A.13}$$

As a result, the wage is determined by value of the marginal product of labor in manufacturing:

$$w^* = P_m A_m.$$

Finally, the value of the marginal product of land in agriculture is equal to the land rental rate:

$$r = P_a MPT_a.$$

2. Consumer maximization: workers and land owners maximize utility described by (A.12) subject to the constraint $P_a c_a + P_m c_m + P_s c_s = y_i$ where $y_L = w$ and $y_T = r$.
3. Services market clearing: $Q_s = C_s$
4. Labor market clearing: $L_m + L_a + L_s = L$.

Employment in agriculture
Note that in an equilibrium with positive manufacturing employment the condition $P_a MPL_a = w^* = P_m A_m$ still defines equilibrium employment in agriculture. Thus, L_a^* is still given by equation (4).

Employment in services
Market clearing for services and consumer utility maximization imply:

$$A_s L_s = Q_s = C_s = \frac{\alpha_s}{P_s}\left(wL + \theta r T\right)$$

The equilibrium relative price of services is $P_s = \frac{w}{A_s}$ as implied by equation (A.13). Thus, equilibrium employment in services is:

$$L_s^* = \alpha_s L + \alpha_s \theta \frac{r^*}{w^*} T. \tag{A.14}$$

A closed form solution for L_s^* can be obtained by multiplying and dividing the second term on the RHS of equation (A.14) by L_a^* and substituting for the solution for L_a^* displayed in equation (4):

$$L_s^* = \alpha_s \left\{ L + \theta \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} \frac{A_T T}{A_L} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{1}{1-\sigma}} \right\}. \tag{A.15}$$

where the equilibrium labor share in agriculture is $\Gamma^* = \gamma^\sigma \left(P_m A_m / P_a A_N A_L\right)^{1-\sigma}$, as in the economy without services. Finally, land rents are:

$$\frac{r^*}{w^*} = \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} A_T \left[\frac{1}{\gamma^\sigma}\left(\frac{P_a A_N}{P_m A_m}\right)^{1-\sigma} - \left(\frac{1}{A_L}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{A.16}$$

Employment in manufacturing can be found by using the labor market clearing condition: $L_m^* =$

$L - L_a^* - L_s^*$.

Technical Change and Structural Transformation

Note that the solution for agricultural employment given by equation (A.5) is independent of the demand structure. This is because wages are set by the value of the marginal product of labor in manufacturing. As a result, the effects of each type of technical change on the agricultural employment share are the same as in the economy without non traded goods, discussed in section C. We call them the supply-side effects of technical change: $\frac{\partial L_a^*}{\partial A_i}$ for $i = N, T, L$.

The effects of technical change on the demand for labor in the service sector are illustrated by equation (A.14). First, note that workers spend a constant share of their labor endowment on services ($\alpha_s L$). This is because the service sector uses only labor for production. Thus, any increase in wages has both an income and substitution effect on the demand for services by workers. The income effect increases their demand for services as their labor endowment is more valuable. The substitution effect reduces the demand for services as their price, the wage, increases. When preferences are Cobb-Douglas both effects have the same magnitude and cancel-out.[4] As a result, technical change can only affect the demand for services through its effect on the consumption of land owners: $\alpha_s \theta \frac{r^*}{w^*} T$. which is an increasing function of land revenues relative to the price of services in terms of labor ($\frac{r^*}{w^*}$). Thus, the effects of agricultural technical change on employment in the service sector depend on its effect on land rents. In turn, agricultural technical change always increases land rents, as we show below. As a result, the demand for services and employment in the service sector increase. We call this the demand side effects of technical change: $\frac{\partial L_s^*}{\partial A_i}$ for $i = N, T, L$.

In what follows, we show that when technical change is Hicks-neutral or land-augmenting, both the supply-side and demand-side effects reduce manufacturing employment. However, when technical change is strongly labor-saving each effect moves manufacturing employment in opposite directions. On the one hand, the supply side effect releases labor from agriculture, increasing the labor supply for manufacturing. On the other hand, the demand-side effect increases labor demand in services, reducing the supply of labor for manufacturing. As a result, the net effect on manufacturing employment depends on the relative strength of each effect. Here, we show that the supply-side effect dominates as long as $\sigma < (1 - \Gamma^*)(1 - \alpha_s \theta)$. Note that because $1 - \alpha_s \theta < 1$, this condition is stronger than the condition required for agricultural technical change to be strongly labor-saving : $\sigma < 1 - \Gamma^*$. Thus, it is satisfied as long as landowners' consumption share of local services ($\alpha_s \theta$) is not too large.

Labor-augmenting technical change

a) Land and labor are strong complements:

a.1) $\frac{\partial L_a^*}{\partial A_L} < 0$, $\frac{\partial L_s^*}{\partial A_L} > 0$, and $\frac{\partial L_m^*}{\partial A_L} > 0$ if $\sigma < (1 - \Gamma^*)(1 - \alpha_s \theta)$

a.2) $\frac{\partial L_a^*}{\partial A_L} < 0$, $\frac{\partial L_s^*}{\partial A_L} > 0$, and $\frac{\partial L_m^*}{\partial A_L} < 0$ if $(1 - \Gamma^*)(1 - \alpha_s \theta) < \sigma < (1 - \Gamma^*)$.

Proof: As shown in section C, when $\sigma < 1 - \Gamma^*$ labor augmenting technical change generates a reallocation of labor away from agriculture. Then, the labor market clearing condition implies that

[4] If instead of Cobb-Douglas we used homothetic preferences with an elasticity of substitution smaller than one, as suggested by Herrendorf, Rogerson and Valentinyi (2013), the income effect would dominate. Thus, the demand for services from workers would be increasing in wages.

$L_m + L_s$ must increase. In addition, $\frac{\partial L_s^*}{\partial A_L} \geq 0$ because land rents increase: $\partial \frac{r^*}{w^*} / \partial A_L > 0$, see equations (A.14) and (A.16). Finally, the effects of technical change on manufacturing employment depend on whether the increase in the demand for labor in the service sector is larger than the reduction in agricultural employment, given by equation (A.9). To make this comparison we obtain $\frac{\partial L_s^*}{\partial A_L}$ from the solution for L_s^* given by equation (A.15), as follows:

$$\begin{aligned}
\frac{\partial L_s^*}{\partial A_L} &= \alpha_s \theta \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} \left\{ -\frac{A_T T}{A_L^2} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{1}{1-\sigma}} + \frac{1}{1-\sigma} \frac{A_T T}{A_L} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{1}{1-\sigma}-1} \frac{\partial \frac{1-\Gamma^*}{\Gamma^*}}{\partial A_L} \right\} \\
&= \alpha_s \theta \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\sigma}{1-\sigma}} \left(\frac{1-\Gamma^*}{\Gamma^*}\right)^{\frac{\sigma}{1-\sigma}} \frac{A_T T}{A_L^2}
\end{aligned}$$

where the last step used the solution for the elasticity of the land to labor share ratio with respect to labor-augmenting technical change given by equation (A.10). Then, $\frac{\partial L_m^*}{\partial A_L} > 0$ as long as $-\frac{\partial L_a^*}{\partial A_L} > \frac{\partial L_s^*}{\partial A_L}$, which requires the following parameter restriction:

$$\sigma < (1-\Gamma^*)(1-\alpha_s\theta). \tag{A.17}$$

b) Land and labor are not strong complements: $\frac{\partial L_a^*}{\partial A_L} > 0$, $\frac{\partial L_s^*}{\partial A_L} \geq 0$, and $\frac{\partial L_m^*}{\partial A_L} < 0$.

Proof: As shown in section C, when $\sigma \geq 1 - \Gamma^*$ labor augmenting technical change generates a reallocation of labor into agriculture. Then, the labor market clearing condition implies that $L_m + L_s$ must fall. In addition, $\frac{\partial L_s^*}{\partial A_L} \geq 0$ because land rents increase: $\partial \frac{r^*}{w^*} / \partial A_L > 0$, see equations (A.14) and (A.16). As a result, manufacturing employment falls.

Land-augmenting technical change: $\frac{\partial L_a^*}{\partial A_T} > 0$, $\frac{\partial L_s^*}{\partial A_T} \geq 0$, and $\frac{\partial L_m^*}{\partial A_T} < 0$.

Proof: See equations (A.5), (A.14), (A.16), and the labor market clearing condition.

Hicks-neutral technical change: $\frac{\partial L_a^*}{\partial A_N} > 0$, $\frac{\partial L_s^*}{\partial A_N} \geq 0$, and $\frac{\partial L_m^*}{\partial A_N} < 0$.

Proof: See equations (A.5), (A.14), (A.16), and the labor market clearing condition.

Alternative assumptions regarding the residence of absentee landowners

In this section we assess to what extent the predictions of the model are affected by alternative assumptions regarding the residence of absentee landowners. Our current treatment of land-owners nests the two standard assumptions in the regional economics literature.[5] The first is that land income accrues to absentee landowners and is thus not spent within the region. This case would correspond to $\theta = 0$. The second is land is publicly owned, or that land income is redistributed lump-sum to workers. Because preferences are homothetic, this case is equivalent to $\theta = 1$. Note that these cases do not take into account the local consumption of landowners who reside in the region under consideration but own land in other regions. Thus, we are implicitly assuming that absentee

[5] For a detailed discussion see Fujita (1989).

landowners reside outside of the set of regions under consideration or the country. Alternatively, we could consider the following two scenarios:

a) All landowners reside within the country, but not necessarily in the region where they own land. In this case, aggregate service consumption in region i would be:

$$C_{i,s} = c_{i,s,L}\ L_i + \sum_j c_{j.s,T}\pi_{ji}T_j$$

where j indexes regions and π_{ji} is the share of landlords owning land in region j but residing in region i. Unfortunately, such detailed data on the residence of landowners is not available. Thus, we make the simplifying assumption that the geographical distribution of landowners' residence is proportional to the distribution of workers's residence: $\pi_{ji} = L_i/L$, where $L = \sum_j L_j$. In this case, aggregate service consumption in region i would be:

$$C_{i,s} = c_{i,s,L}\ L_i + \sum_j c_{j.s,T}\frac{L_i}{L}T_j.$$

Then, consumer utility maximization implies:

$$C_{i,s} = \frac{\alpha_s}{P_{i,s}}\left[w_i L_i + rL_i\right],$$

where r $=(1/L)\sum_j r_j T_j$ are land rents per worker, which take the same value for all regions. In turn, services market clearing implies

$$A_s L_{i,s} = \frac{\alpha_s}{P_{i,s}}\left[w_i L_i + rL_i\right].$$

The equilibrium relative price of services is $P_{i,s} = \frac{w_i}{A_s}$ as implied by equation (A.13), and w_i is set by the tradable manufacturing sector, $w_i = w^* = P_m A_m$.Thus, the equilibrium employment share in services is:

$$\frac{L_{i,s}}{L_i} = \alpha_s + \alpha_s\frac{r}{w^*} \tag{A.18}$$

The first term in equation (A.18) reflects the service consumption of workers. In turn, last term reflects the consumption of absentee landowners, which is spread across all regions in proportion to their worker population. Because this term is identical for all regions, technical change in agriculture in any region increases the employment share of services in the same amount in all regions. Finally, note that, because preferences are homothetic, we would have obtained the same result if we had assumed that land income is taxed by the national government and rebated lump sum to workers.

b) Some landowners reside in the region where they own land, the rest in other regions within the country.

We also consider a combination of the case discussed in the main text of the paper with case a) above by assuming that a share θ_i of landowners reside within the region where they own land and a share $(1-\theta_i)$ resides in other regions within the country. In this case, aggregate service consumption in region i would be:

$$C_{i,s}=c_{i,s,L}\,L_i+c_{i,s,T}\theta_iT_i+\sum_{j\neq i}c_{j.s,T}\left(1-\theta_j\right)\pi_{ji}T_j$$

where j indexes regions and π^{ji} is the share of absentee landlords owning land in region j but residing in region i. We can again make the simplifying assumption that the geographical distribution of absentee landowners is proportional to the distribution of workers within the country: $\pi_{ji}=\frac{L^i}{L}$. In this case aggregate service consumption in region i would be:

$$C_{i,s}=c_{i,s,L}\,L_i+c_{i,s,T}\theta_iT_i+\sum_{j\neq i}c_{j.s,T}\left(1-\theta_j\right)\frac{L_i}{L}T_j$$

Then, consumer utility maximization implies:

$$C_{i,s}=\frac{\alpha_s}{P_{i,s}}\left[w_iL_i+\theta_ir_iT_i+r_{-i}L_i\right],$$

where $r_{-i}\equiv 1/L\sum_{j\neq i}\left(1-\theta_j\right)r_jT_j$ are *absentee* landowner rents per worker. Thus, equilibrium employment share in services is:

$$L_{i,s}=\alpha_sL_i+\alpha_s\theta\frac{r_i}{w^*}T_i+\alpha_s\frac{r_{-i}}{w^*}L_i \tag{A.19}$$

Note that this equation is identical to equation (A.14) up to the second term on the RHS, which reflects the consumption of landowners who both own land and reside within the region. In addition, there is a third term on the RHS which the reflects the consumption of landowners who reside within the region but own land in other regions. However, the effect of agricultural technical change in region i on service employment in region i is the same as in the simpler setup described by equation (A.14). That is, local technical change increases local demand for services to the extent that landowners reside locally, as reflected by the second term of equation (A.19). This is because technical change in region i increases local land rents but does not affect land rents in other regions (see equation (A.16)). More formally, $\partial L^*_{si}/\partial A_{Li}=\left(\theta\alpha_sT_i/w^*\right)\partial r_i/\partial A_{Li}$.

III. Data

This section contains a detailed description of the main variables used in the empirical section of the paper.

Value of output per worker in agriculture. Data on value of agricultural production per worker is from the Brazilian Agricultural Census (IBGE (1996,2006)) and it is sourced from the IBGE online data repository SIDRA.[6]. The variable "value of output per worker in agriculture" is defined as the logarithm of total value of agricultural production divided by the total number of workers employed in agriculture. Total value of production in agriculture in 1996 comes from tables: 500 (value of production by seasonal crop), 513 (value of production by permanent crop), 527 (value of production by horticulture product), 534 (value of production by forestry product), 551 (value of production by vegetable extraction product) and 338 (value of animals by type). Total value of production in agriculture in 2006 comes from tables: 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). Number of workers in agriculture comes from table 321 in 1996 and table 956 in 2006. In Tables A2 and A3 we also use as a control for the initial level of agricultural productivity the value of output per worker in agriculture from the 1985 Agricultural Census. Data for the 1985 Census was manually extracted from the original documentation available from the IBGE online library.[7] Total value of production in agriculture in 1985 comes from table 96 (expenditures, value of production and revenues), while number of workers in agriculture comes from table 87.

The definition of agricultural workers in the Agricultural Censuses of 1996 and 2006 deserves a more detailed discussion. This variable includes: employees, family members employed in farm activities, sharecroppers and people who reside in the farm and perform agricultural activities without a formal contract. Receiving a wage is not a necessary condition to be considered an agricultural worker, so subsistence farmers are also included in this definition. In addition, workers that perform non-agricultural activities within the farm, such as truck drivers, accountants, mechanics or any other worker performing tasks in support of the main activity of the farm are counted as agricultural workers. Instead, family members or farm residents who were not employed in farm activities, domestic workers that are employed within the household of the owner but are not assigned agricultural tasks, and workers that are hired by a service provider company that has a contract with the farm are not counted as agricultural workers. There are two potential issues with this definition of agricultural workers in the Agricultural Census. The first issue is about how seasonal workers are accounted for. In particular, there is a potential double counting problem for seasonal workers that work in more than one farm. This is because the variable is collected at farm level, and not at individual level, and the definition of agricultural workers includes both permanent (employed for at least 180 days during the Census year) and seasonal workers, even when the latter are not employed on the reference date of the Census (p. 53, IBGE, 2006). Therefore, if seasonal workers are employed by more than one farm during a year, and they are recorded by each of them in the

[6]At `www.sidra.ibge.gov.br`.
[7]At `www.biblioteca.ibge.gov.br`

Census, the variable "number of workers" may overestimate the actual number of workers. Notice that seasonal workers account for 10.35% of total workers in 1996 (p. 43, IBGE, 1996) and 14% in 2006 (SIDRA). The second issue is that the variable "number of workers" in the Agricultural Census does not include employees hired by service provider companies that are contracted by the farm to perform agricultural activities. The Agricultural Census of 2006 was the first to report the number of employees of service providers hired by farms (separately from number of workers), while their number was not reported in the 1996 Census. This means that the only way to construct a definition of agricultural workers that is consistent across the Agricultural Censuses of 1996 and 2006 is to exclude workers contracted by service providers in 2006 (IBGE, 2006, p.34). Notice that in 2006, 238.825 agricultural establishments contracted service providers to perform tasks on their land: this accounted for 4.61% of all Brazilian establishments in that year (SIDRA). Notice that these two issues do not affect the agricultural employment variables calculated using the Population Census.

All variables come at the municipality level and have been aggregated at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE.[8] All values are reported at current values. In order to convert 1996 and 2006 values to 2000 Reais we used the Índice Geral de Preços - Disponibilidade Interna (IGP-DI) prepared by the Fundaçao Getulio Vargas. We set the index equal to 1 in 2000 and divided the value of production in 1996 and 2006 by the value of the index in those years. First differences are defined between 1996 and 2006.

Labor intensity in agriculture. Data on labor intensity in agriculture come from the Brazilian Agricultural Census (IBGE (1996,2006)). The variable is defined as the logarithm of total number of workers employed in farms divided by the total area in farms (in hectares). Number of workers in agriculture in 1996 comes from table 321 in 1996 (people employed in farms) and table 956 in 2006 (people employed in farms on 12/31). A detailed discussion of the definition of agricultural workers in the Agricultural Census is reported in the description of the variable "value of output per worker in agriculture". Total land in farms comes from table 314 in 1996 and table 787 in 2006 (both reporting the area in farms). All variables come at the municipality level and have been aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 1996 and 2006.

Employment share in agriculture. Data on employment share in agriculture come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1991, 2000 and 2010). The Supplement reports detailed information on employment for a sample of the Brazilian population that is representative of the whole population at the municipal level. 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.[9] Sector of employment is reported according to the first version of the CNAE-Domiciliar in both 2000 and

[8] At `http://www.ipeadata.gov.br/iframe_dicionario.aspx?width=1074&height=480`

[9] Both in 2000 and in 2010 some workers did not declare a valid sector of employment and were classified as working in a specific category: "*setor maldefinido*"(badly defined sector). These were around 1.2% in 2000 and around 6.2% in 2010. We do not count these workers in the denominator of the employment shares. This is the correct treatment if these workers are drawn randomly from the other sectors of the economy.

2010. We define the agricultural sector as any sector with code from 01000 through 09999 (section "A": *agricultura, pecuária, silvicultura e exploração florestal* and section "B": *pesca*). Notice that this definition includes also workers that are employed by firms that provide services related to agriculture (code 01401 in the CNAE Domiciliar). In the year 1991 sector of employment is reported according to the old classification. For these years, we define the agricultural sector as any sector with code from 010 through 049: these include both agriculture and fishery. In every year, the sector of activity refers to the sector of the firm where the worker is employed, even if the worker performs a task that is not directly related to this activity. Employed people are defined as anyone who reported being employed during the reference week, both in permanent and seasonal jobs. People who during the reference week worked for no compensation helping someone else in the household and people who worked in agriculture or fishery for subsistence are also counted as employed. People who suspended work during the reference week for holidays, strike or leave are also considered as employed. Original data come at the individual level: we aggregate these data at the municipality level using the individual weights provided by the IBGE. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 2000 and 2010. We do not define first differences between 1991 and 2000 because between these two years the IBGE changed its definition of employment in two ways. First, it started to count zero-income workers as employed. Second, the IBGE changed the reference period for considering a person employed: while in 1991 such period included the last 12 months, in 2000 it only included the reference week of the Census. Since these changes made the definition of total employment not homogeneous between 1991 and 2000, we do not compute changes in the employment shares. See text for further details.

Employment share in manufacturing. Data on employment share in manufacturing come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 2000 and 2010). The Supplement reports detailed information on employment for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is calculated as total number of people who reported working in the manufacturing sector divided by the total number of people that reported being employed in any sector of the economy.[10] Sector of employment is reported according to the first version of the CNAE-Domiciliar in both 2000 and 2010. We define the manufacturing sector as any sector with code from 15000 through 37999 (section "D": *indústrias de transformação*). In every year, the sector of activity refers to the sector of the firm where the worker is employed, even if the worker performs a task that is not directly related to this sector. Since 2000, employed people are defined as anyone who reported being employed during the reference week, both in permanent and seasonal jobs. People who during the reference week worked for no compensation helping someone else in the household and people who worked in agriculture or fishery for subsistence are also counted as employed people. People who suspended work during the reference week for holidays, strike or leave are also considered as

[10] Both in 2000 and in 2010 some workers did not declare a valid sector of employment and were classified as working in a specific category: "*setor maldefinido*"(badly defined sector). These were around 1.2% in 2000 and around 6.2% in 2010. We do not count these workers in the denominator of the employment shares.

employed. Original data come at the individual level: we aggregate these data at the municipality level using the individual weights provided by the IBGE. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 2000 and 2010.

Employment in manufacturing. Data on employment in manufacturing come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1991, 2000 and 2010). The Supplement reports detailed information on employment for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is defined as the logarithm of the total number of people who reported working in the manufacturing sector. Sector of employment is reported according to the first version of the CNAE-Domiciliar in both 2000 and 2010. For these years, we define the manufacturing sector as any sector with code from 15000 through 37999 (section "D": *indústrias de transformação*). In the years 1991 and 2000 sector of employment is reported according to the old classfication. For these years, we define the manufacturing sector as any sector with code from 100 through 300. In every year, the sector of activity refers to the sector of the firm where the worker is employed, even if the worker performs an activity that is not directly related to this sector. Employed people are defined as anyone who reported being employed during the reference week. People who suspended work during the reference week for holidays, strike or leave are also considered as employed. Original data come at the individual level: we aggregate these data at the municipality level using the individual weights provided by the IBGE. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 2000 and 2010 and between 1991 and 2000 for the pre-trends regression discussed in section V.B and presented in table A6.

Wage in manufacturing. Data on the wage in manufacturing come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1991, 2000 and 2010). The Supplement reports detailed information on income from employment for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is calculated as the logarithm of the average wage of manufacturing workers 10 years or older in 2000 Reais. See the discussion of employment in manufacturing for details on the definition of the manufacturing sector. The Sample Supplement reports the income from the principal job of every worker. We estimate the average wage in every municipality by looking only to the income from the principal job of workers who report being employees (*empregado com carteira de trabalho assinada* or *empregado sem carteira de trabalho assinada*): this excludes entrepreneurs and self-employed workers, whose income from the principal job may include profits or rents but it includes informal workers.[11] Average wage of manufacturing workers at the municipality level is calculated as the weighted average of the wage of all employees who report working in the manufacturing sector as principal job, using the individual weights provided by the IBGE. Average wage at the level of AMC (micro-region) is calculated as a weighted average of the income in all municipalities belonging to

[11] Income of entrepreneurs and self-employed workers is reported to the Census as business revenues minus business expenses.

an AMC (micro-region) using the number of manufacturing employees in every municipality as weight. All wages are reported in current values of the currency in circulation in the year of the Census: these are thousand Cruzeiros in 1991 and Reais in 2000 and 2010. In order to convert 1991 values to Reais we divided wages in thousand Cruzeiros by 2.75 million. In order to deflate all current values to July 2000 Reais we used the monthly *Índice Nacional de Preços ao Consumidor* (INPC). We set the index equal to 1 in July 2000 (the reference month for the 2000 Census) and divided the 1991 values in Reais by the INPC index in August 1991 (the reference month for the 1991 Census) and the 2010 values by the INPC index in July 2010 (the reference month for the 2010 Census). First differences are defined between 2000 and 2010 and between 1991 and 2000 for the pre-trends regression discussed in section V.B and presented in table A6.

Share of agricultural land cultivated with soy. Data on the share of agricultural land cultivated with soy come from the Brazilian Agricultural Census (IBGE (1996,2006)). The variable is defined as area reaped with soy divided by total land in farms. Area reaped with soy comes from table 501 in 1996 (*área colhida por produtos das lavouras temporárias e condição do produtor*); and table 1823 in 2006 (*produção, venda, valor da produção e área colhida da lavoura temporária por produtos da lavoura temporária e grupos e classes de atividade*). Total land in farms comes from table 314 in 1996 (*área dos estabelecimentos por grupo de atividade econômica e condição legal das terras*) and table 787 in 2006 (*número de estabelecimentos e área dos estabelecimentos agropecuários, por condição legal do produtor em relação ás terras, sexo do produtor, grupos de atividade econômica e grupos de área total*). All variables come at the municipality level and have been aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 1996 and 2006 and in all regressions with this variable we winsorize top and bottom 1% of observations of the variable to reduce the importance of extreme values in our estimates.

Share of agricultural land cultivated with maize. Data on the share of agricultural land cultivated with maize come from the Brazilian Agricultural Census (IBGE (1996,2006)). The variable is defined as area reaped with maize divided by total land in farms. Area reaped with maize comes from table 501 in 1996 (*área colhida por produtos das lavouras temporárias e condição do produtor*); and table 1823 in 2006 (*produção, venda, valor da produção e área colhida da lavoura temporária por produtos da lavoura temporária e grupos e classes de atividade*). Total land in farms comes from table 314 in 1996 (*área dos estabelecimentos por grupo de atividade econômica e condição legal das terras*) and table 787 in 2006 (*número de estabelecimentos e área dos estabelecimentos agropecuários, por condição legal do produtor em relação ás terras, sexo do produtor, grupos de atividade econômica e grupos de área total*). All variables come at the municipality level and have been aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 1996 and 2006 and in all regressions with this variable we winsorize top and bottom 1% of observations of the variable to reduce the importance of extreme values in our estimates.

Share of agricultural land cultivated with genetically engineered soy. Data on the share of agricultural land cultivated with genetically engineered (GE) soy come from the Brazilian Agricultural Census (IBGE (1996,2006)). The variable is defined as area reaped with GE soy divided by total land in farms. Area reaped with GE soy comes from table 824 in 2006 (*produção, venda, valor da produção e área colhida da lavoura temporária por produtos da lavoura temporária, tipo de semente, tipo de colheita, tipo de cultivo e destino da produção*) and it is assumed to be 0 in 1996. Total land in farms comes from table 314 in 1996 (*área dos estabelecimentos por grupo de atividade econômica e condição legal das terras*) and table 787 in 2006 (*número de estabelecimentos e área dos estabelecimentos agropecuários, por condição legal do produtor em relação ás terras, sexo do produtor, grupos de atividade econômica e grupos de área total*). All variables come at the municipality level and have been aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 1996 and 2006 and in all regressions with this variable we winsorize top and bottom 1% of observations of the variable to reduce the importance of extreme values in our estimates.

Share of agricultural land cultivated with traditional soy. Data on the share of agricultural land cultivated with traditional (non-GE) soy come from the Brazilian Agricultural Census (IBGE (1996,2006)). The variable is defined as area reaped with soy minus the area reaped with GE soy divided by total land in farms. Area reaped with soy comes from table 501 in 1996 (*área colhida por produtos das lavouras temporárias e condição do produtor*); and table 1823 in 2006 (*produção, venda, valor da produção e área colhida da lavoura temporária por produtos da lavoura temporária e grupos e classes de atividade*). Area reaped with GE soy comes from table 824 in 2006 (*produção, venda, valor da produção e área colhida da lavoura temporária por produtos da lavoura temporária, tipo de semente, tipo de colheita, tipo de cultivo e destino da produção*) and it is assumed to be 0 in 1996. Total land in farms comes from table 314 in 1996 (área *dos estabelecimentos por grupo de atividade econômica e condição legal das terras*) and table 787 in 2006 (*número de estabelecimentos e área dos estabelecimentos agropecuários, por condição legal do produtor em relação ás terras, sexo do produtor, grupos de atividade econômica e grupos de área total*). All variables come at the municipality level and have been aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE. First differences are defined between 1996 and 2006 and in all regressions with this variable we winsorize top and bottom 1% of observations of the variable to reduce the importance of extreme values in our estimates.

Change in potential soy (maize) yield. Data on the change in potential soy (maize) yield come from the FAO GAEZ v3.0 database.[12] Raw data for potential soy (maize) yield under low and high input levels are reported on Figures A3 and A4 in this appendix (Figures A5 and A6 for maize). To construct the change in potential soy (maize) yield we use two variables from the *Suitability and Potential Yield* Series: *Total production capacity for low input level rain-fed soybean (maize)* and *Total production capacity for high input level rain-fed soybean (maize)*.[13] These series

[12] All data can be downloaded from `www.gaez.fao.org`.

[13] The description of each technology in the FAO-GAEZ dataset documentation is as follows. Low-level inputs/traditional management: "Under the low input, traditional management assumption, the farming system is largely subsistence based and not necessarily market oriented. 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 measures." High-level inputs/advanced management: "Under the high input, advanced management assumption, the farming

are aimed at capturing the *potential production capacity in terms of output density which equals total grid cell production potential divided by grid cell area.* Variables are expressed in tons per hectare. The time is set to the baseline 1961-1990, which means that the FAO-GAEZ agricultural model has been applied considering the average climate of the period 1961-1990. To construct the change in potential soy (maize) yield we subtract the low input level variable from the high input level variable in each municipality. Data has then been aggregated at the at AMC-level using the correspondence proposed by IPEA and IBGE.

Share of rural population. Data on the share of rural population come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1980 and 1991). The Supplement reports information on the area where people live (rural or urban) for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is calculated considering only people 10 years or older, as the total number of people living in rural areas divided by the total number of people. The type of area where a person lives is classified under four categories in 1980 (two categories for cities and villages, and two for rural areas, depending on density) and under eight categories in 1991 (three categories for cities and villages, and four categories for rural areas). In both years IBGE defines the boundaries of the cities using the most recent municipal law on the matter. We define rural population as everybody who does not live within the boundary of a city or village. Original data come at the individual level: we aggregate these data at the municipality level using the individual weights provided by the IBGE. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE.

Income per capita. Data on income per capita come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1980, 1991, 2000 and 2010). The Supplement reports information on income received from any source for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is calculated as the natural logarithm of the average real income of people between 10 and 60 years old earning strictly positive income. The IBGE collects data on income coming from both labor and other sources (including pensions, social programs, rents, capital income etc). We defined income for every person as the sum of income coming from all sources. In order to avoid our results being affected by pensions, we restrict our sample to positive income earners between 10 and 60 years old. Original data come at the individual level: we compute average income at the municipality level using the individual weights provided by the IBGE. Average income at the level of AMC is calculated as a weighted average of the income in all municipalities belonging to an AMC using the number of income earners in every municipality as weight. We matched municipalities to AMC and micro-regions using the correspondence proposed by IPEA and IBGE. All incomes are reported in current values of the currency in circulation in the year of the Census: these are Cruzeiros in 1980, thousand Cruzeiros in 1991, Reais in 2000 and 2010. In order to convert values to Reais, we divided

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."

income in Cruzeiros by 2.75 billion in 1980; and we divided income in thousand Cruzeiros by 2.75 million in 1991. In order to deflate all current values to July 2000 Reais we used the monthly *Índice Nacional de Preços ao Consumidor* (INPC). We set the index equal to 1 in July 2000 (the reference month for the 2000 Census) and divided the 1980 and 1991 values in Reais by the INPC index in August 1980 and August 1991 respectively (these were the reference months for the 1980 and 1991 Censuses).

Population density. Data on population density are constructed from the Brazilian Population Censuses (IBGE, 1980 and 1991) and geo-referenced maps of Brazil prepared by GADM (`http://www.gadm.org/`) using data from IBGE. The Brazilian Censuses report information on the total number of people of any age living in a municipality. Geo-referenced on Brazilian municipalities has information on the area of each municipality. The variable is calculated as the natural logarithm of the total number of people living in a municipality divided by the area of the municipality in square kilometers. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE.

Literacy rate. Data on the literacy rate come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1980 and 1991). The Supplement reports information on literacy for a sample of the Brazilian population that is representative of the whole population at the municipal level. 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. In 1980 IBGE classified people who used to be able to read and write but were currently unable to do so in a separate category: we do not consider these people as literate. Original data come at the individual level: we aggregate these data at the municipality level using the individual weights provided by the IBGE. Data at the municipality level have been in turn aggregated at the at the level of AMC and micro-region using the correspondence proposed by IPEA and IBGE.

Wage in agriculture. Data on the wage in agriculture in 1991 come from Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1991). The Supplement reports detailed information on income from employment for a sample of the Brazilian population that is representative of the whole population at the municipal level. The variable is calculated as the logarithm of the average wage of agricultural workers 10 years or older in 2000 Reais. Sector of employment is reported according to the old classfication. We define the agricultural sector as any sector with code from 010 through 049. See the discussion of employment in agriculture for details on the definition of the agricultural sector in 1991. The Sample Supplement reports the income from the principal job of every worker. We estimate the average wage in every municipality by looking only to the income from the principal job of workers who report being employees (*empregado com carteira de trabalho assinada* or *empregado sem carteira de trabalho assinada*). This definition of workers excludes entrepreneurs and self-employed workers, whose income from the principal job may include profits or rents.[14] The definition however includes all informal workers. Average wage of agricultural workers at the municipality level is calculated as the weighted average of the wage

[14]Income of entrepreneurs and self-employed workers is reported to the Census as business revenues minus business expenses.

of all employees who report working in the agricultural sector as principal job, using the individual weights provided by the IBGE. Average wage at the level of AMC (micro-region) is calculated as a weighted average of the income in all municipalities belonging to an AMC (micro-region) using the number of agricultural employees in every municipality as weight. All wages are reported in current values of the currency in circulation in the year of the Census: these are thousand Cruzeiros in 1991: see the discussion of income per capita for details on the way we deflate 1991 wages to 2000 Reais.

Employment in manufacturing firms with more than 30 employees. Data on employment in manufacturing firms with more than 30 employees come from PIA, *Pesquisa Industrial Anual.* The data is collected yearly by the IBGE and made available for research purposes in the facilities of the IBGE - Rio de Janeiro upon approval of a research proposal. We restrict our analysis to the manufacturing sector as defined by CNAE 1.0 (code 15 to 37) and use micro-data from 1996 to 2006. We construct our measure of employment starting from variable $V0194$, which is defined in the original documentation as: *Total pessoal ocupado em 31/12* or end-of-year number of workers. We sum this variable across all plants located in the same municipality in each year. In order to have a representative sample at municipality level we use only plants that are part of firms with 30 or more employees (which are sampled with probability one).

Average wage in manufacturing firms with more than 30 employees. Data on wages in manufacturing firms with more than 30 employees come from PIA, *Pesquisa Industrial Anual.* The data is collected yearly by the IBGE and made available for research purposes in the facilities of the IBGE - Rio de Janeiro upon approval of a research proposal. We restrict our analysis to the manufacturing sector as defined by CNAE 1.0 (code 15 to 37) and use micro-data from 1996 to 2006. We construct our measure of average wage starting from variable $V0195$, which is defined in the original documentation as: *Salários, retiradas e outras remuneraçãoes* and includes total wage remuneration as well as pension contributions and other worker benefits. The average wage is calculated summing this variable across all plants located in the same municipality in each year and dividing it by the total number of workers in the same plants in the same year. In order to have a representative sample at municipality level we use only plants that are part of firms with 30 or more employees (which are sampled with probability one).

Migration rate. The migration rate is estimated with data from the Sample Supplement of the Brazilian Population Censuses (*questionário da amostra*: IBGE, 1991, 2000 and 2010). To construct net migration in a municipality between two Census years we follow the cohort average method (Shryock, Siegel and Larmon 1980, pp. 630-635). Letting the first Census year be year 0, and the second Census year be year t, this implies computing, for every age $a+t$, the total number of people living in each municipality m in year t, and subtract from this number the total number of people of that age that one would expect given the population of age a living in that municipality in year 0 and given the average survival rate for that age group in Brazil.

Formally, this implies computing the following formula:

$$M_{m,a+t} = \frac{M'_{m,a+t} + M''_{m,a+t}}{2}$$

where $M'_{m,a+t}$ and $M''_{m,a+t}$ are net migrants in municipality m computed with the "forward survival rate method" and with the "reverse survival rate method", respectively:

$$\begin{aligned} M'_{m,a+t} &= P^t_{m,a+t} - s^t_a P^0_{m,a} \\ M''_{m,a+t} &= \frac{P^t_{m,a+t}}{s^t_a} - P^0_{m,a} \end{aligned}$$

In these equations, $P^t_{m,a+t}$ is the total number of people of age $a+t$ living in municipality m in year t and $P^0_{m,a}$ is the total number of people of age a observed in the same municipality in year 0. s^t_a is the probability that a person of age a in year 0 is still alive after t years. Both $P^0_{m,a}$ and $P^t_{m,a+t}$ are observed in the Brazilian Censuses of 1991, 2000 and 2010. Survival probabilities are estimated using data on the whole Brazilian population with the formula:

$$s^t_a = \frac{P^t_{a+t}}{P^0_a}$$

where $P^0_a = \sum_m P^0_{m,a}$ and $P^t_{a+t} = \sum_m P^t_{m,a+t}$. Notice that with this formula net migration for Brazil as a whole is zero by construction.

Let M^t_{a+t} be the estimate of net number of migrants in a municipality with age $a+t$ in Census year t: total net migration M^t is the sum of net migrants with age between 10 and 60 years in the initial period[15]. We compute migration rates MR^t_m as net migrants in a municipality divided by the total number of people living in that municipality in year 0:

$$MR^t_m = \frac{M^t_m}{P^0_m}$$

The cohort average method implicitly imposes two assumptions that are worth discussing briefly. First, the method assumes that Brazil is a closed country (i.e. nobody emigrate or immigrate from abroad between two Census years). Second, the method assumes that in a given year everybody has the the same survival probability conditional on age. We discuss these assumptions in turn.

We deal with the closed country assumption by computing net migration only of people born in Brazil.[16] In other words, in every Census year, we compute $P^0_{m,a}$, $P^t_{m,a+t}$ and s^t_a using only people that report being born in Brazil, which means that our migration rates refer to Brazilian-born only. We do this because people born outside of Brazil are on average more likely to enter or leave the country between two Census years. Note however, that this decision creates three additional issues. First, it requires to interpret the survival probability as the probability of re-observing a person somewhere in Brazil after t years, rather than the probability of this person of simply being alive after t years. Second, it implicitly assumes that nobody born in Brazil emigrated outside of Brazil before year 0 and then returned between the year 0 and the year t. Finally, it entirely abstracts from migration of foreigners (both within the country and from abroad). The first of

[15] Since the last three Censuses were taken in 1991, 2000 and 2010, we can estimate migration between 1991 and 2000 and between 2000 and 2010. In the first case migrants are aged between 19 and 69 years in 2000, while for the migration rate between 2000 and 2010 they are aged between 20 and 70 in 2010.

[16] This is what Shryock, Siegel and Larmon (1980) recommend. The standard alternative is to compute the survival probabilities only of Brazilian-born and then apply these probabilities to both nationals and foreigners.

these three issues is inconsequential, but the other two may bias our estimates if Brazil as a whole experienced large migration flows between 1991 and 2010, or if foreign-born people accounted for a large share of the population. In practice this does not seem to be the case. In 2010, only 0.19% of the Brazilian-born population aged 15 to 70 reported having moved to Brazil from abroad during the previous 10 years. Moreover, people born outside of Brazil always account for less than 1% of the total population living in Brazil between 1991 and 2010 (the exact figures are: 0.52% in 1991, 0.4% in 2000 and 0.31% in 2010).

In order to allow different survival probability across different groups of the population we compute net migration and survival probabilities separately for men and women and then add them to compute total migration.

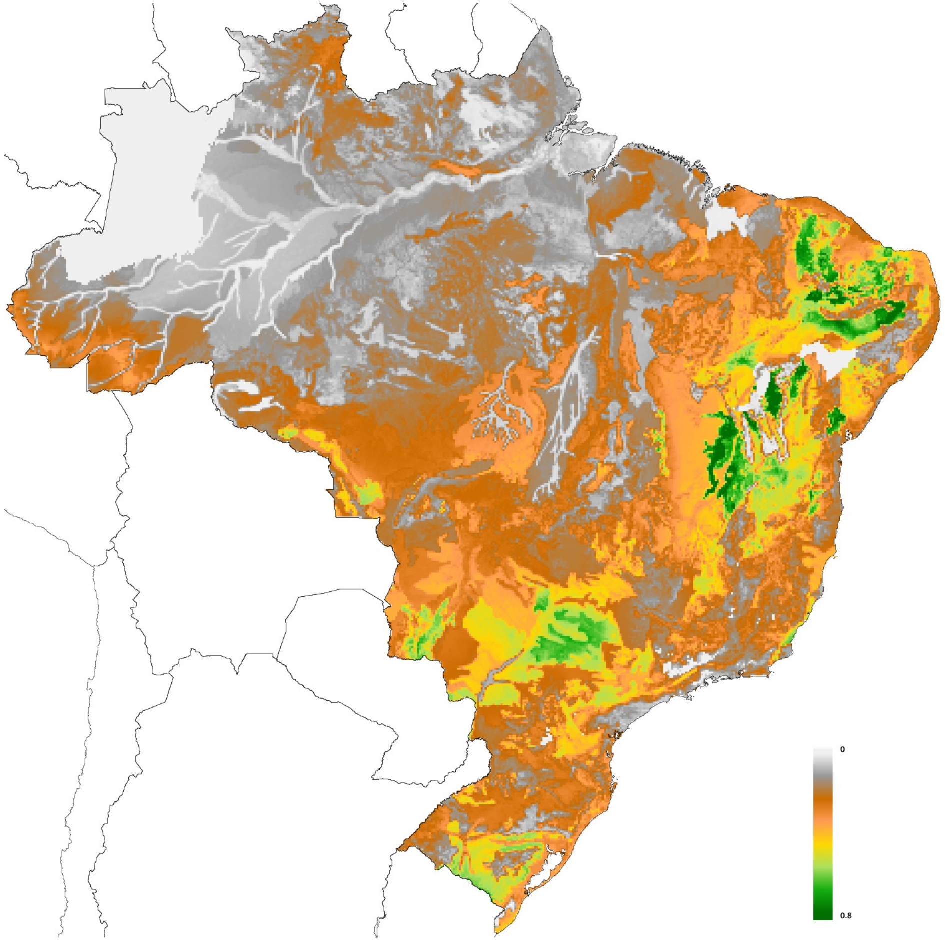


Figure A3 Potential soy yield under low agricultural technology

Notes: Data from FAO-GAEZ.

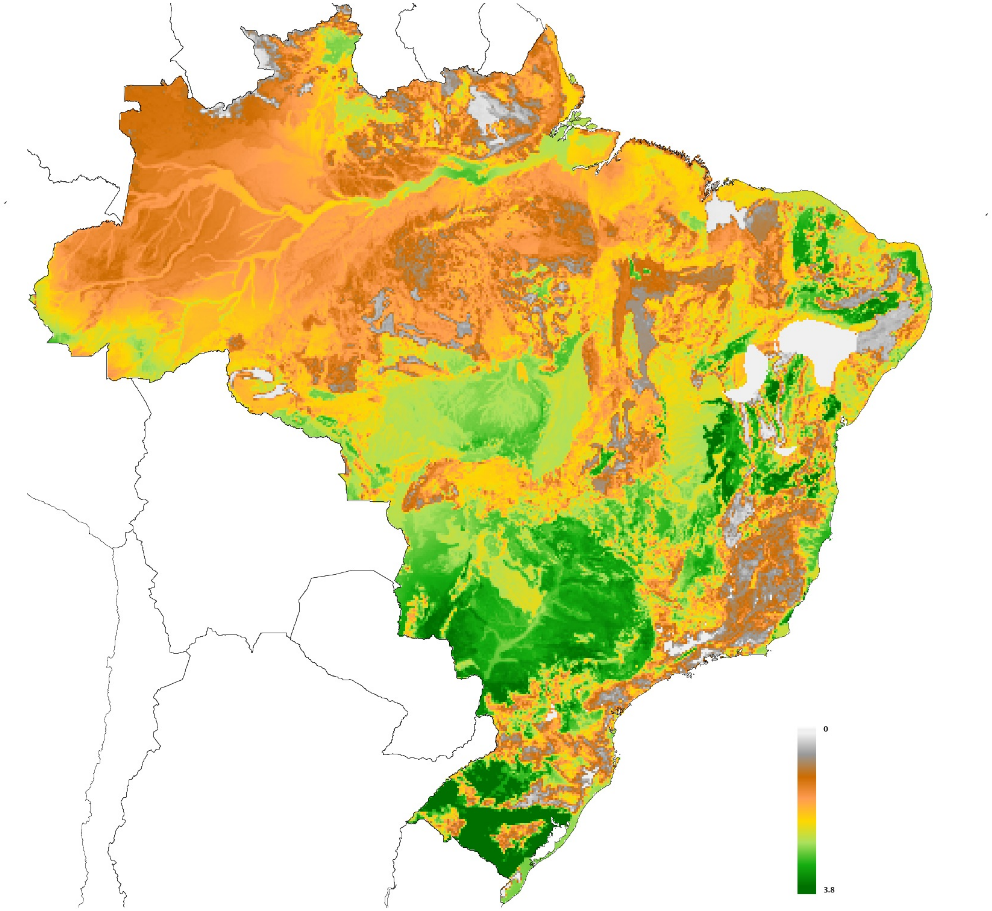


Figure A4 Potential soy yield under high agricultural technology

Notes: Data from FAO-GAEZ.

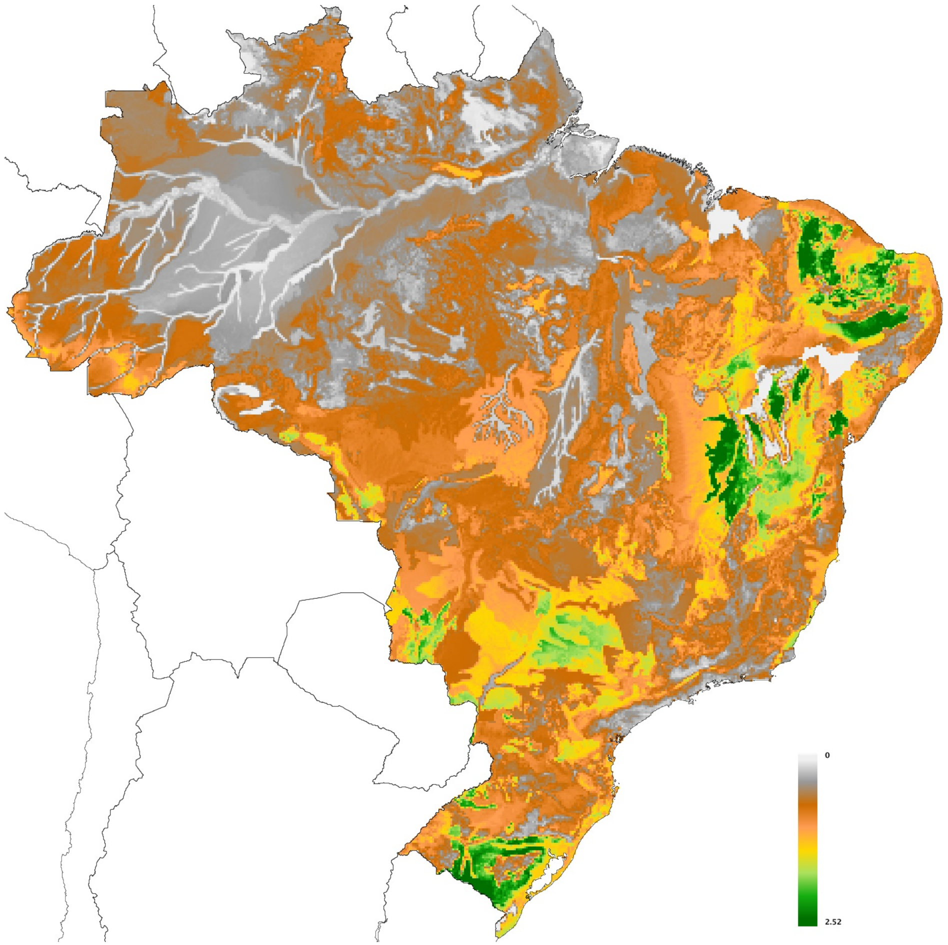


Figure A5 Potential maize yield under low agricultural technology

Notes: Data from FAO-GAEZ.

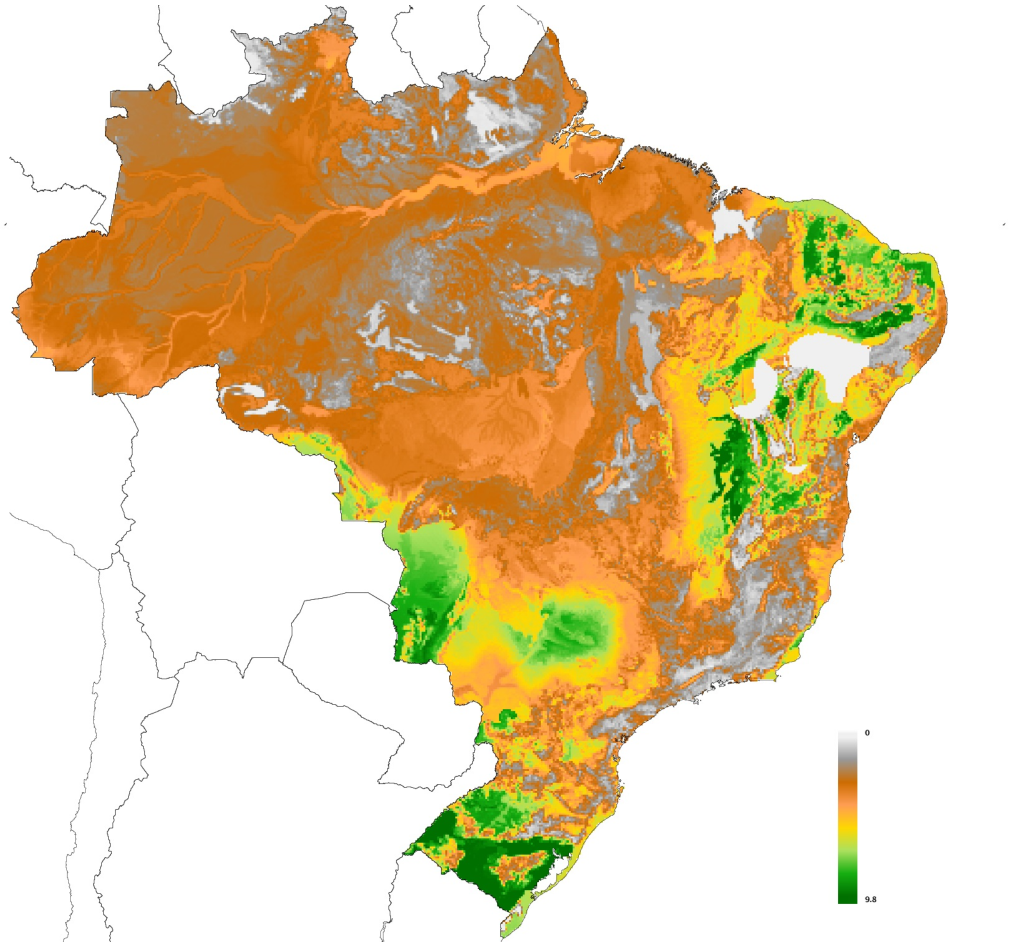


Figure A6 Potential maize yield under high agricultural technology

Notes: Data from FAO-GAEZ.

IV. Empirics

In this section we include additional figures and tables discussed in section IV of the paper.

A. *Basic Correlations in the Data*

No further results required.

B. *Empirical Strategy*

No further results required.

C. *Agricultural Outcomes*

No further results required.

D. *Manufacturing Outcomes*

No further results required.

E. *Services and Other Sectors*

Table A1: The effect of agricultural technological change on employment shares
Observations are weighted by share of aggregate employment

	Δ Log output per worker	Δ Employment share			
	in agriculture	Agriculture	Manufacturing	Services	Other Sectors
	(1)	(2)	(3)	(4)	(5)
ΔA^{soy}	0.205**	-0.011***	0.011***	0.001	-0.000
	(0.088)	(0.002)	(0.002)	(0.002)	(0.001)
ΔA^{maize}	-0.097**	0.003***	-0.001	-0.002*	-0.001
	(0.039)	(0.001)	(0.001)	(0.001)	(0.001)
Share Rural Pop	0.153	-0.078***	-0.006	0.067***	0.017***
	(0.231)	(0.008)	(0.008)	(0.007)	(0.005)
Log Income per Capita	-0.166	0.006**	-0.016***	0.008*	0.001
	(0.205)	(0.003)	(0.005)	(0.005)	(0.002)
Log Pop Density	-0.003	0.003***	-0.004***	0.002***	-0.001**
	(0.036)	(0.001)	(0.001)	(0.001)	(0.000)
Literacy Rate	-0.003	0.027**	0.044***	-0.052***	-0.020***
	(0.529)	(0.014)	(0.014)	(0.015)	(0.008)
Observations	4,149	4,149	4,149	4,149	4,149
R-squared	0.037	0.331	0.118	0.162	0.112

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 when the data sources are the Agricultural Censuses of 1996 and 2006 (column 1), and over the years 2000 and 2010 when the data sources are the Population Censuses of 2000 and 2010 (columns 2-5). All municipality controls are from the Population Census of 1991. Estimation method is weighted least squares, with weights equal to the share of aggregate employment. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Basic Correlations in the data

Income per Capita and Services Employment Share

	Δ Income per capita (1)	Δ Employment share in services (2)	Δ Income per capita (3)	Δ Employment share in services (4)
Δ Soy area share × Agri Fam ^{soy}			0.994***	0.134*
			(0.322)	(0.069)
Δ Soy area share	0.120	0.007	-0.540**	-0.105**
	(0.074)	(0.019)	(0.226)	(0.050)
Δ Maize area share	0.080*	0.033**	0.074*	0.033**
	(0.043)	(0.015)	(0.043)	(0.015)
Agri Fam ^{soy}			0.005	0.008**
			(0.012)	(0.003)
Observations	3,765	3,765	3,765	3,765
R-squared	0.002	0.002	0.008	0.004

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). Changes in explanatory variables are calculated over the years 1996 and 2006 (source: Agricultural Censuses). Agri Fam^{soy} is the share of farms whose main activity in 2006 was soy production and that were classified as family farms. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A3: The Effect of Agricultural Technical Change on Income per Capita and Services Employment Share

	Δ Income per capita	Δ Employment share in services	Δ Income per capita	Δ Employment share in services
	(1)	(2)	(3)	(4)
Δ A ^{soy}× Agri Fam ^{soy}			0.019*	-0.001
			(0.010)	(0.003)
ΔA^{soy}	0.018***	-0.002	0.013**	-0.002
	(0.005)	(0.002)	(0.006)	(0.002)
ΔA^{maize}	0.001	-0.000	0.000	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
Agri Fam ^{soy}			0.002	0.019***
			(0.023)	(0.006)
Share Rural Pop	0.037**	0.043***	0.018	0.034***
	(0.017)	(0.005)	(0.017)	(0.006)
Log Income per Capita	-0.110***	-0.015***	-0.110***	-0.015***
	(0.010)	(0.003)	(0.010)	(0.003)
Log Pop Density	0.010***	0.000	0.009***	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Literacy Rate	0.110***	-0.009	0.079**	-0.022**
	(0.030)	(0.010)	(0.031)	(0.010)
Observations	4,149	4,149	4,149	4,149
R-squared	0.093	0.103	0.098	0.109

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). Changes in explanatory variables are calculated over the years 1996 and 2006 (source: Agricultural Censuses). All municipality controls are from the Population Census of 1991. Agri Fam^{soy} is the share of farms whose main activity in 2006 was soy production and that were classified as family farms. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

F. *Variable Factor Endowments*

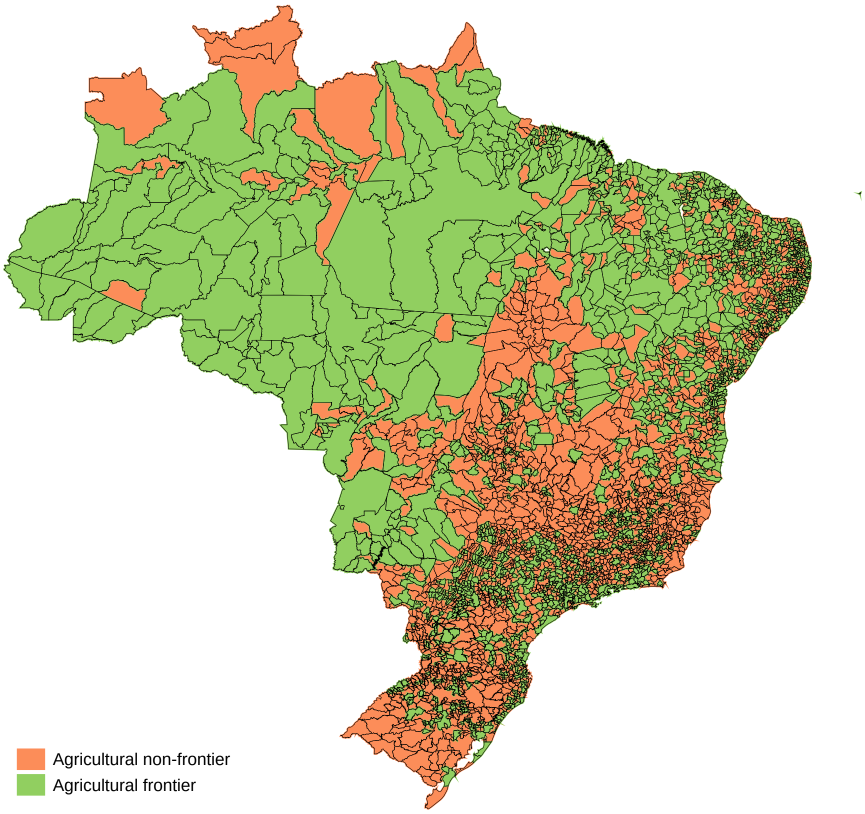


Figure A7 Agricultural Frontier
Municipalities

Notes: Data from Brazilian Agricultural Censuses of 1996 and 2006.

V. Robustness Checks

In this section we include additional figures and tables discussed in section V of the paper.

Table A4: The effect of technological change on agriculture

Robustness to controlling for additional initial municipality characteristics

PANEL A	Δ Soy area			Δ Maize area		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔA^{soy}	0.012***	0.013***	0.011***	-0.000	0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
ΔA^{maize}	-0.000	-0.001	-0.000	0.003***	0.003**	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Agri Labor Prod	0.001			-0.001		
	(0.001)			(0.002)		
Log Avg Agri Wage		-0.003**			-0.008***	
		(0.002)			(0.003)	
Agri Employment Share			0.023***			0.012
			(0.006)			(0.009)
Baseline controls	Y	Y	Y	Y	Y	Y
Observations	3,841	3,683	3,841	4,032	3,872	4,032
R-squared	0.118	0.117	0.121	0.015	0.015	0.015

PANEL B	Δ Log output per worker			Δ Log labor intensity			Δ Employment share		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔA^{soy}	0.172***	0.132***	0.128***	-0.092***	-0.061***	-0.069***	-0.016***	-0.021***	-0.011***
	(0.029)	(0.028)	(0.028)	(0.022)	(0.021)	(0.021)	(0.003)	(0.002)	(0.002)
ΔA^{maize}	-0.045***	-0.034***	-0.031**	0.042***	0.031***	0.035***	0.004***	0.006***	0.003***
	(0.013)	(0.012)	(0.012)	(0.009)	(0.009)	(0.009)	(0.001)	(0.001)	(0.001)
Log Agri Labor Prod	-0.063***			0.048***			-0.008***		
	(0.022)			(0.018)			(0.002)		
Log Avg Agri Wage		-0.012			-0.041			-0.003	
		(0.036)			(0.028)			(0.003)	
Agri Employment Share			0.113			0.095			-0.164***
			(0.127)			(0.090)			(0.011)
Baseline controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	4,149	3,985	4,149	4,149	3,985	4,149	4,149	3,985	4,149
R-squared	0.014	0.013	0.012	0.010	0.008	0.007	0.079	0.071	0.128

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 when the data sources are the Agricultural Censuses of 1996 and 2006 (PANEL A: all columns, PANEL B: columns 1 to 6), and over the years 2000 and 2010 when the data sources are the Population Censuses of 2000 and 2010 (PANEL B: columns 7 to 9). In Panel B: Log Agricultural Labor Productivity is from the 1985 Agricultural Census in columns 1 to 6 and from the 1996 Agricultural Census in columns 7 to 9. All other municipality controls are from the Population Census of 1991. Baseline controls include: share of rural population, log income per capita, log population density and literacy rate. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A5: The effect of technological change on manufacturing
Robustness to controlling for additional initial municipality characteristics

	Δ Employment share			Δ Log employment			Δ Log wage		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔA^{soy}	0.013***	0.022***	0.014***	0.110***	0.186***	0.155***	-0.026**	-0.018	-0.034***
	(0.002)	(0.002)	(0.002)	(0.021)	(0.020)	(0.020)	(0.013)	(0.013)	(0.013)
ΔA^{maize}	-0.002**	-0.004***	-0.002**	-0.021**	-0.042***	-0.034***	0.015***	0.012**	0.017***
	(0.001)	(0.001)	(0.001)	(0.009)	(0.009)	(0.009)	(0.006)	(0.005)	(0.005)
Log Agri Labor Prod	0.013***			0.124***			0.004		
	(0.001)			(0.013)			(0.008)		
Log Avg Agri Wage		-0.008***			-0.016			-0.029*	
		(0.002)			(0.023)			(0.016)	
Agri Employment Share			0.114***			0.535***			0.164***
			(0.009)			(0.087)			(0.054)
Baseline controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	4,149	3,985	4,149	4,149	3,985	4,149	4,149	3,985	4,149
R-squared	0.102	0.077	0.118	0.090	0.070	0.077	0.045	0.045	0.047

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). Log Agricultural Labor Productivity is from the 1996 Agricultural Census. All other municipality controls are from the Population Census of 1991. Baseline controls include: share of rural population, log income per capita, log population density and literacy rate. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

B. *Pre-Existing Trends*

Table A6: The effect of agricultural technological change on manufacturing and migration

Manufacturing employment, manufacturing wages and net migration

Robustness to controlling for pre-existing trends

	Δ Log employment $_t$ (1)	Δ Log wage $_t$ (2)	Migration rate $_t$ (3)
ΔA^{soy} × After $_t$	0.234***	-0.117***	-0.015**
	(0.025)	(0.018)	(0.006)
ΔA^{maize} × After $_t$	-0.060***	0.053***	0.010***
	(0.012)	(0.008)	(0.003)
ΔA^{soy}	0.007	0.062***	-0.004
	(0.020)	(0.014)	(0.005)
ΔA^{maize}	-0.004	-0.027***	-0.003
	(0.010)	(0.006)	(0.002)
Share Rural Pop $_{t-2}$	0.255***	0.010	-0.128***
	(0.042)	(0.027)	(0.014)
Log Income per Capita $_{t-2}$	0.011	-0.058***	0.045***
	(0.027)	(0.020)	(0.006)
Log Pop Density $_{t-2}$	-0.016***	-0.003	-0.006**
	(0.006)	(0.004)	(0.002)
Literacy Rate $_{t-2}$	0.245***	0.190***	0.017
	(0.081)	(0.052)	(0.019)
After $_t$	-0.233***	0.138***	-0.020***
	(0.032)	(0.022)	(0.006)
Observations	7,984	7,984	7,984
R-squared	0.031	0.018	0.096

Notes: Dependent variables calculated using data from the Population Censuses of 1991, 2000 and 2010. Municipality controls are from the Population Censuses of 1980 and 1991. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

C. *Larger Unit of Observation: Micro-Regions*

Table A7: The effect of agricultural technological change on manufacturing
Employment share, employment and wages
Robustness to using a larger unit of observation: micro-regions

	Δ Employment share (1)	Δ Log employment (2)	Δ Log wage (3)
ΔA^{soy}	0.017***	0.139***	-0.022
	(0.004)	(0.029)	(0.016)
ΔA^{maize}	-0.003	-0.037***	0.016**
	(0.002)	(0.013)	(0.007)
Share Rural Pop	0.014	0.017	-0.103
	(0.012)	(0.121)	(0.089)
Log Income per Capita	-0.002	0.058	-0.168**
	(0.007)	(0.088)	(0.073)
Log Pop Density	0.004***	0.030***	-0.032***
	(0.001)	(0.011)	(0.007)
Literacy Rate	0.016	0.007	0.128
	(0.021)	(0.261)	(0.180)
Observations	557	557	557
R-squared	0.101	0.107	0.239

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). All municipality controls are from the Population Census of 1991. The unit of observation is the micro-region. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

D. *Input-Output Linkages*

Table A8: The effect of agricultural technological change on manufacturing
Employment share, employment and wages
Robustness to excluding sectors directly linked to soy and maize

	Δ Employment share (1)	Δ Log employment (2)	Δ Log wage (3)
ΔA^{soy}	0.013***	0.167***	-0.011
	(0.002)	(0.021)	(0.016)
ΔA^{maize}	-0.004***	-0.057***	0.010
	(0.001)	(0.009)	(0.007)
Share Rural Pop	0.012***	0.042	-0.014
	(0.004)	(0.058)	(0.044)
Log Income per Capita	-0.002	0.075*	-0.117***
	(0.002)	(0.038)	(0.027)
Log Pop Density	0.003***	0.034***	-0.040***
	(0.000)	(0.008)	(0.006)
Literacy Rate	0.025***	0.086	0.144*
	(0.007)	(0.124)	(0.084)
Observations	4,149	4,134	4,059
R-squared	0.037	0.042	0.030

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). All municipality controls are from the Population Census of 1991. Manufacturing sectors directly linked to soy and maize are: food and beverages (code 15), manufacturing of other chemicals (code 24090) and manufacturing of goods from refined oil (code 23020). The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The number of observations in columns 2 and 3 is smaller because in some municipalities sectors directly linked to soy and maize account for the whole manufacturing sector.

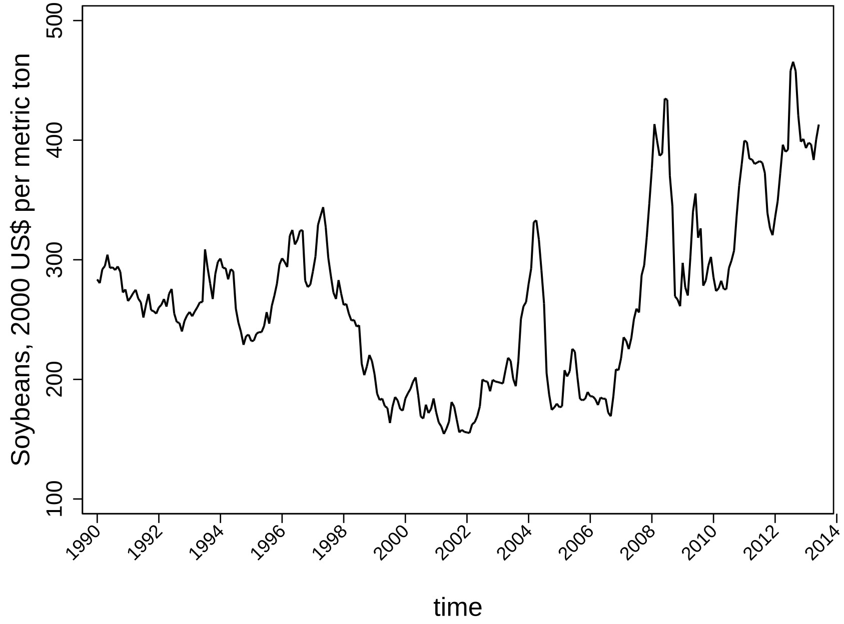


Figure A8 Evolution of soy price (1990-2013)

Notes: The Figure shows the monthly evolution of soy real price between 1990 and 2013. Data are from the IMF Primary Commodity Prices database, series code: *PSOYB_USD*, expressed in nominal US$ per metric ton. We deflate the series using the US *Consumer Price Index for All Urban Consumers: All Items*, source: Federal Reserve St. Louis, series code: *CPIAUCNS*, rescaled so that 2000 is the base year.

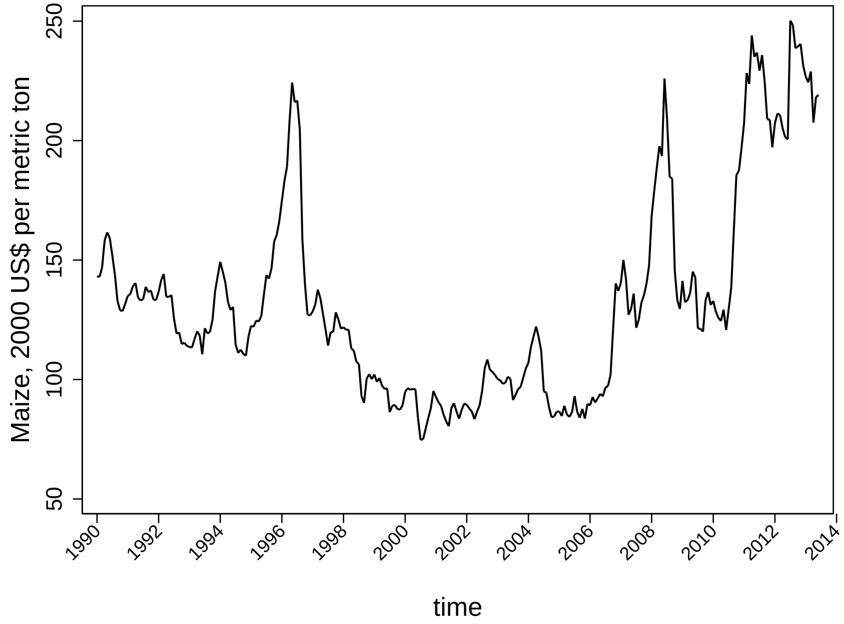


Figure A9 Evolution of maize price (1990-2013)

Notes: The Figure shows the monthly evolution of maize real price between 1990 and 2013. Data are from the IMF Primary Commodity Prices database, series code: *PMAIZMT_US*, expressed in nominal US$ per metric ton. We deflate the series using the US *Consumer Price Index for All Urban Consumers: All Items*, source: Federal Reserve St. Louis, series code: *CPIAUCNS*, rescaled so that 2000 is the base year.

Table A9: The effect of agricultural technological change on manufacturing

Data on employment and wages from the Annual Manufacturing Survey (PIA)

Robustness of results reported in Table 9 to controlling for commodity prices

	Log Total Employment $_t$			Log Wage $_t$		
	(1)	(2)	(3)	(4)	(5)	(6)
A^{soy}	0.122***	0.122***	0.096***	-0.026*	-0.026*	-0.018
	(0.030)	(0.030)	(0.029)	(0.013)	(0.013)	(0.013)
A^{maize}	-0.031**	-0.031**	-0.027*	0.013**	0.013**	0.011*
	(0.015)	(0.015)	(0.015)	(0.006)	(0.006)	(0.006)
Share Rural Pop $\times t$	0.052***	0.052***	0.019	0.030***	0.030***	0.025***
	(0.012)	(0.012)	(0.016)	(0.005)	(0.005)	(0.007)
$P^{soy}A^{soy}$		-0.001	-0.000		0.000	0.000
		(0.001)	(0.001)		(0.001)	(0.001)
$P^{maize}A^{maize}$		-0.001	-0.001		0.000	0.000
		(0.001)	(0.001)		(0.000)	(0.000)
Literacy Rate $\times t$			-0.098***			0.014
			(0.037)			(0.017)
Log Pop Density $\times t$			-0.010***			0.001
			(0.002)			(0.001)
Log Income per Capita $\times t$			0.019*			-0.009**
			(0.010)			(0.004)
Observations	25,262	25,262	25,262	25,239	25,239	25,239
R-squared	0.923	0.923	0.923	0.778	0.778	0.778

Notes: A^{soy} is defined as potential soy yield under high inputs for the years between 2003 and 2007, and the potential soy yield under low inputs for the years between 1996 and 2002. A^{maize} is defined as potential maize yield under high inputs for the years between 2003 and 2007, and potential maize yield under low inputs for the years between 1996 and 2002. P^zA^z *controls* stand for the interaction of the potential yield of soy and maize under low inputs interacted with price levels of these crops between 1996 and 2007. The unit of observation is the municipality. Standard errors clustered at AMC level reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

F. *Spatial Correlation*

Table A10: The effect of technological change on agriculture

Soy and maize expansion

Robustness of results reported in Table 7 to correcting standard errors for spatial correlation

	Δ Soy area share		Δ Maize area share	
	(1)	(2)	(3)	(4)
ΔA^{soy}	0.013	0.013		0.001
Robust standard errors	(0.001)***	(0.002)***		(0.003)
Microregion-clustered standard errors	(0.002)***	(0.003)***		(0.004)
Mesoregion-clustered standard errors	(0.004)***	(0.005)**		(0.005)
Conley standard errors: cutoff 50 Km	(0.002)***	(0.003)***		(0.004)
Conley standard errors: cutoff 100 Km	(0.002)***	(0.004)***		(0.005)
Conley standard errors: cutoff 200 Km	(0.003)***	(0.006)**		(0.006)
ΔA^{maize}		-0.001	0.003	0.003
Robust standard errors		(0.001)	(0.001)***	(0.001)***
Microregion-clustered standard errors		(0.001)	(0.001)***	(0.002)*
Mesoregion-clustered standard errors		(0.003)	(0.002)*	(0.002)
Conley standard errors: cutoff 50 Km		(0.001)	(0.001)***	(0.002)*
Conley standard errors: cutoff 100 Km		(0.002)	(0.001)***	(0.002)
Conley standard errors: cutoff 200 Km		(0.002)	(0.001)**	(0.002)
Rural Pop Share	Y	Y	Y	Y
Controls	N	Y	N	Y
Observations	3,652	3,652	3,652	3,652
R-squared	0.067	0.124	0.009	0.015

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 (source: Agricultural Censuses). Controls include: log income per capita, log population density and literacy rate. The unit of observation is the municipality. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A11: The effect of technological change on agriculture

Productivity, labor intensity and employment share

Robustness of results reported in Table 8 to correcting standard errors for spatial correlation

	Δ Log output per worker (1)	Δ Log labor intensity (2)	Δ Employment share (3)
ΔA^{soy}	0.134	-0.064	-0.021
Robust standard errors	(0.027)***	(0.021)***	(0.002)***
Microregion-clustered standard errors	(0.033)***	(0.026)**	(0.004)***
Mesoregion-clustered standard errors	(0.034)***	(0.030)**	(0.006)***
Conley standard errors: cutoff 50 Km	(0.032)***	(0.024)***	(0.003)***
Conley standard errors: cutoff 100 Km	(0.037)***	(0.028)**	(0.005)***
Conley standard errors: cutoff 200 Km	(0.039)***	(0.032)**	(0.006)***
ΔA^{maize}	-0.033	0.033	0.006
Robust standard errors	(0.012)***	(0.009)***	(0.001)***
Microregion-clustered standard errors	(0.016)**	(0.012)***	(0.002)***
Mesoregion-clustered standard errors	(0.016)**	(0.015)**	(0.003)*
Conley standard errors: cutoff 50 Km	(0.015)**	(0.011)***	(0.002)***
Conley standard errors: cutoff 100 Km	(0.016)**	(0.013)**	(0.002)***
Conley standard errors: cutoff 200 Km	(0.017)*	(0.015)**	(0.003)**
Controls	Y	Y	Y
Observations	4,149	4,149	4,149
R-squared	0.012	0.007	0.073

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 when the data sources are the Agricultural Censuses of 1996 and 2006 (columns 1 and 2), and over the years 2000 and 2010 when the data sources are the Population Censuses of 2000 and 2010 (column 3). Controls include: log income per capita, log population density and literacy rate. The unit of observation is the municipality. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A12: The effect of technological change on manufacturing

Employment share, employment and wages

Robustness of results reported in Table 9 to correcting standard errors for spatial correlation

	Δ Employment share (1)	Δ Log employment (2)	Δ Log wage (3)
ΔA^{soy}	0.021	0.186	-0.024
Robust standard errors	(0.002)***	(0.020)***	(0.012)*
Microregion-clustered standard errors	(0.004)***	(0.030)***	(0.014)*
Mesoregion-clustered standard errors	(0.006)***	(0.045)***	(0.019)
Conley standard errors: cutoff 50 Km	(0.003)***	(0.024)***	(0.014)*
Conley standard errors: cutoff 100 Km	(0.004)***	(0.030)***	(0.015)
Conley standard errors: cutoff 200 Km	(0.006)***	(0.037)***	(0.018)
ΔA^{maize}	-0.004	-0.043	0.014
Robust standard errors	(0.001)***	(0.009)***	(0.005)**
Microregion-clustered standard errors	(0.002)**	(0.014)***	(0.006)**
Mesoregion-clustered standard errors	(0.003)	(0.027)	(0.008)*
Conley standard errors: cutoff 50 Km	(0.001)***	(0.011)***	(0.006)**
Conley standard errors: cutoff 100 Km	(0.002)**	(0.015)***	(0.006)**
Conley standard errors: cutoff 200 Km	(0.003)	(0.020)**	(0.007)*
Controls	Y	Y	Y
Observations	4,149	4,149	4,149
R-squared	0.073	0.068	0.045

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). Controls include: log income per capita, log population density and literacy rate. The unit of observation is the municipality. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

G. *Alternative Definition of Technical Change*

Table A13: The effect of technological change on agriculture and manufacturing
Robustness to Alternative Definition of Technical Change

	Agricultural outcomes					Manufacturing outcomes		
	Δ Soy area share	Δ Maize area share	Δ Log output per worker	Δ Log labor intensity	Δ Employment share	Δ Employment share	Δ Log empl.	Δ Log wage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ A $^{soy\ (h-m)}$	0.020***	0.002	0.164***	-0.145***	-0.019***	0.014***	0.155***	-0.040**
	(0.003)	(0.004)	(0.038)	(0.029)	(0.003)	(0.003)	(0.028)	(0.016)
Δ A $^{maize\ (h-m)}$	-0.001	0.004**	-0.041**	0.063***	0.005***	-0.001	-0.030**	0.023***
	(0.001)	(0.002)	(0.018)	(0.013)	(0.002)	(0.001)	(0.014)	(0.008)
Share Rural Pop	0.038***	0.010	0.098	-0.179***	-0.069***	0.002	-0.017	-0.010
	(0.005)	(0.007)	(0.073)	(0.050)	(0.007)	(0.005)	(0.056)	(0.034)
Log Income per Capita	-0.000	-0.005	-0.012	0.040	0.012***	0.004	0.111***	-0.106***
	(0.002)	(0.004)	(0.048)	(0.038)	(0.004)	(0.003)	(0.037)	(0.026)
Log Pop Density	-0.001***	0.004***	-0.017	-0.019*	0.000	0.001	0.015*	-0.035***
	(0.001)	(0.001)	(0.012)	(0.010)	(0.001)	(0.001)	(0.008)	(0.005)
Literacy Rate	0.056***	-0.009	-0.342**	-0.088	-0.007	0.030***	0.150	0.100
	(0.006)	(0.012)	(0.148)	(0.116)	(0.014)	(0.010)	(0.119)	(0.075)
Observations	3,652	3,652	4,149	4,149	4,149	4,149	4,149	4,149
R-squared	0.150	0.015	0.011	0.011	0.064	0.050	0.056	0.045

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 when the data sources are the Agricultural Censuses of 1996 and 2006 (columns 1 to 4), and over the years 2000 and 2010 when the data sources are the Population Censuses of 2000 and 2010 (columns 5 to 8). $\Delta A^{soy\ (h-m)}$ is defined as potential soy yield under high inputs minus potential soy yield under intermediate level of inputs. $\Delta A^{maize\ (h-m)}$ is defined as potential maize yield under high inputs minus potential maize yield under intermediate level of inputs. Changes calculated over the years 1996 and 2006 when the data sources are the Agricultural Censuses of 1996 and 2006, and over the years 2000 and 2010 when the data sources are the Population Censuses of 2000 and 2010. The unit of observation is the municipality. Robust standard errors reported in parentheses. Significance levels: $p<0.01$, $p<0.05$, $p<0.1$.

References

CONAB. 2003-2012. "Levantamento De Avaliação Da Safra." *Brasília, Brazil: Companhia Nacional de Abastecimento (CONAB)*, http://www.conab.gov.br/conteudos.php?a=1252&t= (accessed April 10, 2013).

Fujita, Masahisa. 1989. *Urban Economic Theory: Land Use and City Size.* Cambridge university press.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2013. "Two Perspectives on Preferences and Structural Transformation." *American Economic Review*, 103(7): 2752–2789.

IBGE. 1980; 1991; 2000; 2010. "Censo Demográfico." *Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE)*, http://loja.ibge.gov.br/ (accessed March 1, 2013).

IBGE. 1981-2011. "Pesquisa Nacional de Amostra de Domicilios (PNAD)." *Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE)*, http://ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2014/default.shtm (accessed July 18, 2013).

IBGE. 1985; 1995-6; 2006. "Censo Agropecuário." *Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE)*, http://www.sidra.ibge.gov.br/bda/acervo/acervo9.asp?e=c&p=CA&z=t&o=3 (accessed June 10, 2012).

IBGE. 1996-2007. "Pesquisa Industrial Anual - Empresa." *Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE).*

IBGE. 2003. *Metodologia do Censo Demográfico.* Vol. 25, Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE).

IMF. 1990-2013. "IMF Primary Commodity Prices database." *Washington, DC: International Monetary Fund (IMF)*, http://www.imf.org/external/np/res/commod/index.aspx (accessed November 16, 2013).

Shryock, Henry S., Jacob S. Siegel, and Elizabeth A. Larmon. 1980. *The methods and materials of demography.* Vol. 2, Department of Commerce, Bureau of the Census.

US Federal Reserve. 1990-2013. "US Consumer Price Index for All Urban Consumers: All Items." *St. Louis, MO: Federal Reserve Bank*, https://research.stlouisfed.org/fred2/series/CPIAUCSL (accessed November 16, 2013).