Industrialization without Innovation

Paula Bustos
CEMFI and CEPR

Juan Manuel Castro Vincenzi
Princeton

Joan Monras
UPF, CREI, BGSE, and CEPR

Jacopo Ponticelli*
Northwestern, NBER and CEPR

First draft: August 2017. This draft: August 2020

Abstract

The introduction of labor-saving technologies in agriculture can release workers who find occupation in the manufacturing sector. The traditional view is that this structural transformation process leads to economic growth. However, if workers leaving agriculture are unskilled, the labor reallocation process reinforces comparative advantage in the least skill-intensive manufacturing industries. We embed this mechanism in a multi-sector endogenous growth model where only skill-intensive manufacturing industries innovate and generate knowledge spillovers. In this setup, the increase in the relative size of the unskilled-labor intensive industries reduces the incentives to innovate and slows down growth. We test the predictions of the model in the context of a large and exogenous increase in agricultural productivity in Brazil. We use social security data to develop a new measure of the labor input in innovation which is representative at any level of spatial aggregation. We find that regions adopting the new agricultural technology experienced a reallocation of unskilled workers away from agriculture into the least R&D-intensive manufacturing industries. The expansion of low-R&D industries attracted workers away from innovative occupations in high-R&D industries, slowing down local aggregate manufacturing productivity growth.

Keywords: Agricultural Productivity, Skill-Biased Technical Change, Innovation, Labor Mobility, Genetically Engineered Soy, Brazil.

JEL codes: F16, J43, O13, O14, O33, O41.

*Bustos: CEMFI and CEPR, paula.bustos@cemfi.es. Castro Vincenzi: Princeton, castro.vincenzi@princeton.edu. Monras: UPF, CREI, Barcelona GSE and CEPR, jm3364@gmail.com. Ponticelli: Northwestern University, NBER and CEPR, jacopo.ponticelli@kellogg.northwestern.edu. We received valuable comments from Thibault Fally (discussant), Manuel García-Santana (discussant), Donald Davis, Gene Grossman, Michael Peters, Diego Puga, Andres Rodriguez-Clare, Chris Tonetti, Chris Udry, Jose P. Vasquez, and seminar participants at CEMFI, CREI, University of Lugano, Columbia, Northwestern, Bank of Spain, the NBER Economic Consequences of Trade, McGill University and the Princeton Growth conferences. We are grateful to acknowledge financial support from the European Research Council Starting Grant 716338. Mark He provided excellent research assistance.

1 Introduction

Early development economists noted that the reallocation of workers from agriculture to manufacturing was fundamental to sustain long run growth (Lewis 1954, Kuznets 1973). This structural transformation process can lead to higher output because labor productivity is lower in agriculture than in the rest of the economy (Caselli 2005, Restuccia, Yang, and Zhu 2008, Lagakos and Waugh 2013). In addition, the manufacturing sector is characterized by economies of scale and knowledge spillovers. As a result, industrialization can lead to higher long run growth (Krugman 1987, Lucas 1988, Matsuyama 1992a). In this paper we qualify these views by noting that manufacturing productivity growth depends not only on the size of the industrial sector but also on its composition (Grossman and Helpman 1991). Thus, if workers leaving the agricultural sector are mostly unskilled, the structural transformation process can reinforce comparative advantage in non-innovating industries, reducing long run growth.

We study the effects of structural transformation on industrial development in the context of a large and exogenous increase in agricultural productivity due to the adoption of genetically engineered (GE) soy in Brazil. This new technology requires fewer but relatively high-skilled workers to yield the same output, thus can be characterized as unskilled-labor-saving technical change. In addition, the technology had heterogeneous effects on yields across areas with different soil and weather characteristics, which permits to estimate the effect of local agricultural technical change on local structural transformation.

To guide our empirical work, we build a three sector endogenous growth model and analyze the free trade equilibrium for each region within a country. The agricultural sector produces an homogeneous traded good using land, skilled and unskilled labor. The manufacturing sector has two industries, H and L. The H industry is skilled-labor intensive and uses non-traded differentiated intermediate inputs for production. The expansion in the number of input varieties used enhances overall productivity in this industry as it facilitates labor specialization as in Adam Smith (1776), Ethier (1982) and Romer (1987). In contrast, new intermediate inputs do not increase productivity in the L industry. This is because this is a traditional, unskilled-labor intensive industry where the scope for process innovation is lower. Still, innovations in the local H-industry have positive productivity spillovers on the local L-industry.[1] Intermediate inputs are produced by monopolistically competitive firms which use their profits to invest in R&D and invent new input varieties. In equilibrium, profits from introducing new varieties are proportional to demand, which

[1] For example, suppose that the L industry is food processing and the H industry is the computer industry. The introduction of a new food additive does not necessarily increase labor productivity in the food industry. However, the introduction of new programs to organize production in the local H industry can generate new management practices that improve the division of labor in the H industry and spill over to the local L industry making it more productive.

is given by the size of each industry. Thus, the growth rate of output in the regional economy is determined by the relative size of the H industry.

In this setup, we model the introduction of GE soy seeds as a skilled-labor-augmenting technical change in agriculture. We show that when skilled and unskilled workers are imperfect substitutes and land and labor are strong complements in production, this type of technical change leads to a reduction in the marginal product of unskilled labor in agriculture. As a result, there is a reduction in labor demand in agriculture and an excess supply of unskilled workers. In equilibrium, unskilled workers reallocate towards the manufacturing sector, reinforcing comparative advantage in the L industry.[2] The larger size of the L industry increases incentives to invest in the development of intermediate inputs for this industry relative to the H industry. As a result, the H industry conducts less R&D, generates less local knowledge spillovers and aggregate regional output grows at a slower pace.

In our empirical work, we attempt to isolate the mechanisms highlighted in the model by tracing the effects of the agricultural productivity shock generated by the introduction of GE soy seeds from the agricultural sector to the rest of the economy. First, to identify the effects of this new technology we exploit variation in the increase in potential soy yields across regions of Brazil as in Bustos, Caprettini, and Ponticelli (2016). This measure of technical change in soy production is a function of weather and soil characteristics, not of actual yields. As a result, it permits to assess the causal effects of agricultural technical change on industrial specialization, innovation and growth by comparing the evolution of variables of interest across micro-regions differently exposed to the new technology.[3]

We start by tracing the flow of workers with different education levels across sectors using detailed individual information from the decadal Brazilian Population Census. We find that the adoption of GE soy led to a reallocation of unskilled workers away from agriculture and towards the manufacturing sector.[4] Our estimates indicate that micro-regions with a one standard deviation higher increase in potential soy yields experienced a 2.4 percentage points larger decrease in the share of unskilled workers employed in agriculture, and a corresponding 2.1 percentage points larger increase in the share of unskilled workers employed in manufacturing. We confirm these findings using yearly formal employment data from Social Security Records (RAIS) which, in addition, shows that the labor reallocation process starts right after the introduction of GE seeds.

[2]This result requires that the L manufacturing industry is not much more skill intensive than agriculture. Otherwise, unskilled workers are absorbed again by the agricultural sector. This is because of Hecksher-Ohlin forces: an increase in the relative supply of a factor generates an expansion of the sector using that factor intensively (Rybczinsky Theorem).

[3]Our geographical unit of observation are Brazilian micro-regions. Micro-regions consist of a group of municipalities and can be thought of as small open economies that trade in agricultural and manufacturing goods but where production factors are immobile.

[4]We classify skilled workers as those who completed the 8th grade, which is equivalent to graduating from middle school in the US.

Next, we study the consequences of the reallocation of unskilled labor from agriculture to manufacturing for industrial specialization. From the point of view of the manufacturing sector, the reallocation of unskilled workers amounts to an increase in the relative supply of unskilled labor. We document, using Population Census data, that this inflow of unskilled workers was completely absorbed by an expansion of the manufacturing industries in the lowest quartile of skill-intensity. A key implication of the model is that specialization in low-skill intensive industries slows down manufacturing productivity growth. This is because as low-skill-intensive industries expand, the return to introduce new intermediate inputs in these industries increases relative to the high-skill industries. Then, investment in product development reallocates to low-skill industries. However, these are traditional industries with lower scope for generating productivity enhancing innovations. In addition, they do not generate knowledge spillovers towards other industries. As a result, productivity growth slows down both in low and high skill industries. To test this mechanism we need to identify which are the industries with the highest and lowest scope for innovation and assess the effect of agricultural technical change on their size, innovation investment and productivity growth. In what follows we discuss how we perform each of these three steps.

First, we source industry-level measures of expenditures in research and development (R&D) from the Industrial Innovation Survey (PINTEC) to show that the manufacturing industries that expanded are in the lowest quartile of expenditure in research and development as a share of sales. This is consistent with their characterization in the model as traditional industries with low scope for developing productivity enhancing innovations. However, standard innovation surveys such as PINTEC do not permit to assess the consequences of this change in industrial specialization for innovation because they are based on a sample of firms which is not representative at fine levels of spatial aggregation.[5] To overcome this problem, we propose a new measure of investment in innovation which is representative at any level of geographical aggregation as it can be constructed using social security data, which covers the universe of formal firms, as described below.

We construct a new measure of investment in innovation activities based on textual analysis of the task descriptions of more than 2500 occupations in RAIS. Tasks generating innovations include, for example, developing new products and processes, creating prototypes, or optimizing methods of production. We use this measure to document that, in regions more exposed to agricultural technical change, the inflow of low-skill agricultural workers into low-R&D manufacturing industries was followed by a reallocation of innovation workers away from high-R&D industries. In particular, micro-regions with a one standard deviation larger increase in potential soy yields experienced a 20 percent larger

[5] Alternative measures of innovation such as patents might be geographically representative but are not representative of the type of innovations which are most frequent in developing countries. According to PINTEC, only 20% of firms which introduced innovations in the period 1997-2008 filed a patent application.

decline in innovation expenditures in high-R&D industries, measured as the wage bill of workers in innovative occupations.

Finally, we test the predictions of the model regarding the effects of agricultural technical change on manufacturing productivity growth. For this purpose, we use data from the yearly manufacturing survey (PIA) which allows us to observe the evolution of labor productivity in the manufacturing sector. We find that micro-regions facing faster agricultural technical change experienced a slowdown in manufacturing productivity growth. Our estimates imply that micro-regions with a one standard deviation larger increase in potential soy yields experienced a 14.8 percent larger increase in the relative size of the low-skill intensive industry and a corresponding 1.2 percent lower yearly growth rate of manufacturing productivity. This decrease in manufacturing productivity is not simply due to a composition effect. As predicted by the model, it is driven by a reduction in productivity growth within both high- and low-R&D intensive industries.

Overall, our empirical findings indicate that unskilled-labor-saving technical change in agriculture can lead to a reallocation of workers towards unskilled-labor-intensive manufacturing industries. This leads to an expansion of the industrial sectors with the lowest R&D intensity in the economy, decreasing innovation in high R&D industries and aggregate manufacturing productivity. We interpret this result as a cautionary tale on the effects of structural change on aggregate productivity growth. The adoption of new technologies in agriculture may result in static productivity gains in the agricultural sector but dynamic losses in manufacturing productivity.

Our findings suggest that different forces driving structural transformation can lead to different types of industrial specialization. In most countries, the process of labor reallocation from agriculture to manufacturing can be ascribed to one of two forces: "push" forces, such as new agricultural technologies that push workers out of agriculture, or "pull" forces, such as industrial productivity growth, that pull workers into manufacturing. We show that when labor reallocation from agriculture to manufacturing is driven by unskilled-labor-saving technical change in agriculture – rather than manufacturing productivity growth – it can generate an expansion in those manufacturing sectors with the lowest potential contribution to aggregate productivity. In this sense, our results are informative for low- to middle-income countries where a large share of the labor force is employed in agriculture, and who import new agricultural technologies from more developed countries.

Related Literature

There is a long tradition in economics of studying the links between agricultural productivity and industrial development. Nurkse (1953), Schultz (1953), and Rostow (1960) argued that agricultural productivity growth was an essential precondition for the industrial revolution. Classical models of structural transformation formalized their ideas by proposing two main mechanisms through which agricultural productivity can speed up industrial growth in closed economies. First, agricultural productivity growth increases income, which can increase the relative demand for manufacturing goods, driving labor away from agriculture and into manufacturing (see Murphy, Shleifer, and Vishny 1989, Kongsamut, Rebelo, and Xie 2001, Gollin, Parente, and Rogerson 2002, Boppart 2014). Second, if productivity growth in agriculture is faster than in manufacturing and these goods are complements in consumption, the relative demand for agricultural goods does not grow as fast as productivity and labor reallocates toward manufacturing (Baumol 1967, Ngai and Pissarides 2007).[6] Note that these two mechanisms are not operative in open economies, where high agricultural productivity induces a reallocation of labor towards agriculture, the comparative advantage sector (Matsuyama 1992b). However, Bustos et al. (2016) show that, if agricultural technical change is labor-saving, increases in agricultural productivity can lead to a reallocation of labor towards the industrial sector, even in open economies, depending on whether land and labor are strong complements which is the focus of their empirical investigation.

Several scholars argue that reallocating agricultural workers into manufacturing can increase aggregate productivity.[7] First, there might be large static productivity gains when labor reallocates from agriculture to manufacturing. Sizable productivity and wage gaps between agriculture and manufacturing have been measured in several studies and have been shown to be larger in developing economies (e.g., Caselli 2005, Restuccia et al. 2008, Lagakos and Waugh 2013, Gollin, Lagakos, and Waugh 2014). To the extent that these gaps arise from the existence of inefficiencies and frictions in the economy, a reallocation of labor from agriculture to the other sectors of the economy is both productivity- and welfare-enhancing.[8] Second, there can be dynamic productivity gains when labor reallocates towards manufacturing if this sector is subject to agglomeration externalities and knowledge spillovers (Krugman 1987, Lucas 1988, Matsuyama 1992a).[9]

In this paper, we take a different perspective based on endogenous growth theory, which stresses that manufacturing productivity growth not only depends on the size of the industrial sector, but also on its composition. In particular, we build on the work of Grossman and Helpman (1991) who study open economy endogenous growth models. In

[6]See also: Caselli and Coleman 2001, Acemoglu and Guerrieri 2008, Buera, Kaboski, and Rogerson 2015.

[7]Although this view has been recently challenged by Franck and Galor (2019) who argue, in line with this paper, that the type of industrial specialization is what determines long-run growth.

[8]More recently, Herrendorf and Schoellman (2018) measure and compare agricultural wage gaps in countries in different stages of the structural transformation process. They find that the implied barriers to labor reallocation from agriculture are smaller than usually thought in the macro-development literature, and argue that labor heterogeneity and selection are important drivers of such gaps. Other scholars emphasize that structural change can be growth-enhancing or growth-reducing depending on the correlation between changes in employment shares and productivity levels (McMillan and Rodrik (2011) and McMillan, Rodrik, and Sepulveda (2017)).

[9]Recent evidence suggests that this channel may be operative in some circumstances. Peters (2019) uses the displacement of Eastern Germans towards Western Germany to show that places experiencing larger population growth specialized in manufacturing and saw GDP per capita grow over the long run.

their model there are two manufacturing industries with different skill intensities that use differentiated intermediates with the same intensity. As a result, incentives for inventing new goods depend on the opportunity cost of performing R&D, which is driven by the skill premium. In contrast, in our model the incentive to do R&D depends on the relative size of the two industries, as in Romer (1990). As a result, an increase in the supply of unskilled labor generates an expansion of the unskilled-labor intensive industry and a reduction in the growth rate. Note that this is not the case in Grossman and Helpman (1991), where an expansion of the supply of unskilled workers does not affect the growth rate. This is so because if both industries are active in the trade equilibrium, there is factor price equalization and, hence, an increase in the supply of unskilled workers does not affect the skill premium, the opportunity cost of innovation, thus the growth rate remains constant.

This paper also builds upon the literature studying the effects of agricultural technical change, particularly those papers that provide evidence that technological advancements in agriculture are skill-biased. For instance Foster and Rosenzweig (1996), who study the effects of the introduction of high-yield varieties in India, show that technological innovations in agriculture increased the relative demand for skill in agriculture and thus returns to primary schooling.[10] We contribute to this literature by showing that the recent introduction of GE soy was also skill-biased. More importantly, we study the implications of skill-biased agricultural technical change for industrialization, which have not previously been explored.

In terms of empirical findings, our paper is also related to contemporaneous work by Imbert, Seror, Zhang, and Zylberberg (2019), who exploit short-run agricultural shocks in China to document how migration from rural to urban areas reduces labor costs and makes firms expand labor usage. They find that firms reduce capital-biased technology adoption in response to these labor supply shocks. Differently from Imbert et al. (2019) we focus on technology adoption in agriculture as the factor driving structural transformation. In terms of outcomes, we do not focus on the capital intensity of manufacturing technology but we study the effects of the reallocation of unskilled workers on industrial specialization through a Hecksher-Ohlin comparative advantage mechanism and its relationship with endogenous growth forces such as innovation investments and their impact on manufacturing productivity dynamics.

The rest of the paper is organized as follows. Section 2 describes the theoretical framework that guides our empirical investigation. Section 3 describes the institutional background, the data, our identification strategy and the empirical results. Finally, section 4 contains our final remarks.

[10]In related recent work, Bragança (2014) shows that investments in soybean adaptation in Central Brazil in the 1970s induced positive selection of labor in agriculture.

2 Model

In this section we describe the model that guides our empirical work. Our model gives rise to a number of predictions on the effects of agricultural technical change on structural transformation, industrial specialization, and economic growth that are useful to interpret the evidence that we present in Section 3. We provide further details of the model and formal proofs of all the results in Appendix C. The model highlights the mechanisms that explain how a positive agricultural shock (i.e. GE soy adoption) may result in a (temporary) slow down in manufacturing productivity growth.

2.1 General setting

The model describes a small region which is open to goods trade. The region produces one agricultural good and two manufacturing goods using land, skilled and unskilled labor. We assume that these production factors are perfectly mobile across sectors but immobile across regions. In what follows, we describe consumer preferences and production technologies in each sector.

Preferences

This economy is populated by infinitely lived consumers that maximize life-time utility. We assume that consumers have constant relative risk aversion flow utility given by:

$$u(c) = \frac{c^{1-\eta} - 1}{1 - \eta}, \tag{1}$$

where c is the composite of consumption of the three goods in the economy: one agricultural good, and two manufacturing goods.[11] Life-time utility is given by $\int e^{-\rho t} u(c(t))dt$, where ρ is the discount factor and t indexes (continuous) time. The budget constraint of the representative consumer is given by $p(t)\vec{c}(t) + \dot{a}(t) \leq w(t) + ra(t)$, where $p(t)$ is the vector of prices and $\vec{c}(t)$ is the vector of consumption quantities. $a(t)$ denotes savings and $w(t)$ wages. In what follows we omit time t when it does not lead to confusion. We assume no asset trade across regions nor with the rest of the world.[12]

Agriculture

The agricultural sector produces an homogeneous final good combining labor and land in a constant elasticity of substitution (CES) production function. In turn, labor is a CES

[11] For simplicity we assume that c is a CES composite. However, given our assumption of a small open economy, demand and supply within each period only determine trade patterns, as long as the three sectors of activity are active.

[12] Alternatively, we can assume open capital markets but binding borrowing constraints so that local returns to investment can be above the international interest rate.

composite of high- and low-skilled labor. The agricultural production function is defined by:

$$Y_a = K_t^h Q_a = K_t^h[(A_L L_a)^{\frac{\sigma-1}{\sigma}} + (A_T T_a)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where A_L and A_T are labor-augmenting and land-augmenting technologies, respectively, and σ is the elasticity of substitution between labor (L_a) and land (T_a). K_t^h is a Hicks-neutral technology parameter reflecting the level of knowledge in the local economy at time t. We discuss this term in detail below. In turn, L_a is a CES aggregate of high- and low-skilled labor:

$$L_a = [(A_U U_a)^{\frac{\varepsilon-1}{\varepsilon}} + (A_S S_a)^{\frac{\varepsilon-1}{\varepsilon}}]^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε is the elasticity of substitution between the two labor types. A_U and A_S are unskilled- and skilled-labor augmenting technologies, respectively.

Manufacturing sector

The manufacturing sector has two *industries* which produce traded homogeneous final goods using high- and low-skilled labor, albeit with different intensities. In addition, production requires the use of non-traded differentiated intermediate inputs.

H-industry The first industry, which we call the H-industry, uses labor and intermediates to produce an homogeneous good using the following technology:

$$Y_m^h = K_t^h Q_m^h = F_m^h(U_m^h, S_m^h)^\alpha (\int^{K_t^h} (x_k^h)^{1-\alpha} dk), \tag{3}$$

where $F_m^h(.)$ is a skill-intensive technology which allows to combine skilled and unskilled labor into one aggregate labor input which we call L_m^h; x_k^h is the quantity used of a given intermediate input k and K_t^h is the total amount of input varieties in the industry. Note that this production function implies that the expansion of input variety enhances overall productivity in the industry. To see this, assume that each intermediate is produced in the same amount $x = X_t/K_t$, as it will be the case in equilibrium. Then, output is increasing in K_t^h given factor inputs U, S and X_t. Similar production functions have been used by Romer (1987) and Ethier (1982) to illustrate a situation in which a larger variety of inputs or machines gives rise to higher total factor productivity as it facilitates labor specialization in the spirit of the pin factory example by Adam Smith (1776).

L-industry The second industry, which we call the L-industry, uses labor and intermediates to produce an homogeneous good using the following technology:

$$Y_m^\ell = K_t^h Q_m^\ell = K_t^h \left(\frac{F_m^\ell(U_m^\ell, S_m^\ell)}{K_t^\ell}\right)^\alpha \left(\int^{K_t^\ell} (x_j^\ell)^{1-\alpha} dj\right) \tag{4}$$

Note that the expansion of the total amount of intermediates in this industry (K_t^ℓ) does not lead to higher productivity. This is because this is a traditional industry where the scope for process innovation is lower. As a result, new intermediate inputs do not facilitate the division of labor nor increase productivity. Still, innovations in the H-industry have positive productivity spillovers on the L-industry. For example, suppose that the L-industry is food processing and the H-industry is the computer industry. The introduction of a new food additive does not necessarily increase labor productivity in the food industry. However, the introduction of new programs to organize production in the local H-industry can generate new management practices that spill over to the local L-industry making it more productive. Note that the model refers to non-traded intermediate inputs, hence these production process innovations are specific to the needs of the industry in a given region in Brazil and only generate knowledge spillovers to other industries within the region.

Intermediate inputs are non-traded differentiated goods produced by monopolistically competitive firms. Each firm is the owner of a blueprint to produce one differentiated intermediate good using one unit of the final good of the industry.

Innovation and growth New intermediate goods are produced by competitive R&D firms using a research technology which invests one unit of the final good of the targeted industry to produce a new intermediate with success probability η. Then, the measure of varieties in the H industry K_t^h grows at a rate that is proportional to R&D investment (I_h):

$$\dot{K}_t^h = \eta I_h.$$

With a similar equation for the L industry as both use the same research technology.

2.2 Equilibrium

In this section we describe the equilibrium conditions of the regional economy. We start by discussing the optimal behavior of intermediate goods producers and R&D firms in the manufacturing sector. Next, we describe the equilibrium in final goods and factor markets.

Intermediate goods producers

Intermediate goods producers optimally choose production to maximize profits which, in the case of the H-industry, are given by:

$$\Pi^h_k = p^h_k x^h_k - x^h_k \tag{5}$$

Note that the demand of inputs from final good producers is given by the marginal product of each input in the final good production:

$$p^h_k = \frac{\partial Y^h_m}{\partial x_k} = (1-\alpha) F^h_m(U^h_m, S^h_m)^\alpha x_k^{-\alpha}$$

We can use this price in equation (5) to find the optimal output of each variety of intermediates:

$$x_k = (1-\alpha)^{2/\alpha} F^h_m(U^h_m, S^h_m).$$

And the equilibrium price $p^h_k = p^h = (1-\alpha)^{-1}$. Note that equilibrium output of each variety of intermediates is proportional to $F^h_m(U^h_m, S^h_m)$ which is the aggregate labor input in the H-industry, in efficiency units. We can use this solution to obtain output in the final good industry H:

$$Q^h_m = K^h_t \kappa F^h_m(U^h_m, S^h_m) \tag{6}$$

where $\kappa = (1-\alpha)^{2*(1-\alpha)/\alpha}$. Equilibrium profits of intermediate good producers are given by:

$$\Pi^h_k = \Pi^h = \chi F^h_m(U^h_m, S^h_m)$$

where $\chi = [(1-\alpha)^{(2-\alpha)/\alpha} - (1-\alpha)^{2/\alpha}]$. Hence, in the H-industry we obtain that output of each intermediate good, total output, and profits of intermediate good producers are all proportional to $F^h_m(U^h_m, S^h_m)$, the aggregate labor input in the H-industry.

Similar derivations imply that profits for intermediate good producers in the L-industry are given by:

$$\Pi^\ell_k = \Pi^\ell = \chi \frac{K^h_t p^\ell_m F^\ell_m(U^\ell_m, S^\ell_m)}{K^l_t}.$$

And the price of intermediate inputs in the L industry is $p^\ell_k = p^\ell = (1-\alpha)^{-1}$.

Innovation

R&D firms invent new input varieties and sell the blueprints to intermediate good producers who are willing to pay up to the present value of their profits. As a result, the return from inventing new varieties for the H industry is given by:

$$r^h = \eta \Pi^h = \eta \chi F^h_m(U^h_m, S^h_m).$$

Note that r^h does not decline with the number of intermediates being used. This feature of the model generates endogenous growth, as we will see below. In addition, r^h is proportional to aggregate employment in the H industry $L^h_m = F^h_m(U^h_m, S^h_m)$.

We assume that research technologies for inventing new varieties for the L and H industry are identical. However, in the L-industry profits for inventing new varieties are declining in the stock of varieties available, K^ℓ_t:

$$r^\ell = \eta\Pi^\ell = \frac{\eta\chi K^h_t p^\ell_m F^\ell_m(U^\ell_m, S^\ell_m)}{K^\ell_t}.$$

There is free entry into R&D activities and firms decide whether to invest in innovation in the L or H industry by comparing their returns. Then, the return to innovation is $r = \max\{r^l, r^h\}$, which determines the equilibrium interest rate. For example, if the economy starts with a high number of varieties in the H industry, then r_ℓ is relatively high and there is innovation only in the L industry. As time passes, K^ℓ_t grows until the return of inventing varieties for the L industry is lower or equal than that for the H industry. At this point, innovation takes place in the H industry forever as returns are constant. Note that, in addition, innovation still takes place in the L industry as spillovers from the H industry keep r^l constant as long as varieties grow at the same rate in both industries.

Factor prices

International final goods prices determine factor prices as in the standard Hecksher-Ohlin model. This is because once equilibrium intermediate good production is taken into account, manufacturing industries behave as a constant returns to scale sector. To see this, note that the production technology for the H industry can be described by equation 6. This implies that we can write the zero profit conditions in the three final good industries using unit cost functions, as follows:

$$p_a = c_a(w_s, w_u, w_T, A_s)/K^h_t \tag{7}$$

$$1 = c^h_m(w_s, w_u, p^h, K^h_t) = (c^h_m(w_s, w_u, 1)^{1-\alpha}(p^h)^\alpha)/K^h_t \propto c^h_m(w_s, w_u, 1)/K^h_t \tag{8}$$

$$p^\ell_m = c^\ell_m(w_s, w_u, p^\ell, K^h_t) = (c^\ell_m(w_s, w_u, 1)^{1-\alpha}(p^\ell)^\alpha)/K^h_t \propto c^\ell_m(w_s, w_u, 1)/K^h_t \tag{9}$$

To obtain these equations we have used the fact that knowledge spillovers generate Hicks-neutral productivity shocks in agriculture and low-skill manufacturing. In addition, given symmetry and optimal behavior in the intermediate goods market, all intermediates are priced at the same level ($p^h = p^\ell = (1-\alpha)^{-1}$) in each industry and produced in the same quantity x^j, and hence, K^h_t also enters as a Hicks-neutral term in the production of high-skilled manufacturing.

The zero profit conditions imply "conditional" factor price equalization for labor inputs. This is the case because given final and intermediate goods prices, equations 8 and 9 define a system of two equations and two unknowns (w_s, w_u). In turn, land prices are determined by the zero profit condition in the agricultural sector given international prices and the equilibrium prices of labor.

Aggregate output

We define the gross domestic output of the economy as total output minus inputs:

$$GDP = p_a Y_a + p_m^\ell \left(Y_m^\ell - \int^{K_t^h} x_k^\ell \right) + \left(Y_m^h - \int^{K_t^h} x_k^h \right) \tag{10}$$

and the long-run growth rate of the economy as $g = \frac{\dot{GDP}}{GDP}$, where the dot indicates the derivative with respect to time, and where total output in each sector is defined in Equations 2, 3, and 4.

Equilibrium

In this context we define the equilibrium in the economy as:

Definition 1. *Given intra-temporal and inter-temporal consumer preferences given by equation 1, a three sector economy with production functions given by equations 2, 3, and 4, and profits from the production of intermediate goods given by equation 5, we say that the economy is in equilibrium if:*

1. *Given world prices for each sector $\{p_a, 1, p_m^\ell\}$, the representative perfectly competitive firm in each final goods sector maximizes profits.*
2. *Given the demand for intermediates generated in the H- and L-industries, intermediate producers maximize profits by choosing the optimal quantity produced and intermediate goods prices.*
3. *R&D firms decide how many new varieties to invent for each manufacturing industry, which determines the return to R&D and the interest rate.*
4. *Consumers optimally decide how much to consume of each good and how much to save for future consumption.*
5. *Land and labor markets of high- and low-skilled workers clear.*

In what follows, we will investigate how an exogenous change in A_S, i.e. a technology that makes high-skilled agricultural workers more productive, changes the allocation of workers across sectors and how this in turn affects the economy's growth rate.

2.3 Structural transformation

We start the discussion of the model's predictions by investigating how agricultural skill-biased technical change affects the demand for high- and low-skilled workers in agriculture. We first discuss the *relative* demand for high- and low-skilled workers and later the *absolute* demand for low-skilled labor in agriculture.

Proposition 1. *Skilled-labor augmenting technical change in agriculture, represented by A_S, leads to an increase in the relative demand for high skilled workers in agriculture if and only if the elasticity of substitution between high- and low-skilled workers is greater than one ($\varepsilon > 1$).*

Proof. See Appendix C. □

This result essentially follows from Acemoglu (2002). When it is relatively easy to substitute low- for high-skilled labor, when the latter becomes more productive firms want to hire relatively more skilled labor.

Note that, at the same time, this increase in A_S makes the aggregate effective labor input in agriculture L_a increase, which is akin to labor-augmenting technical change in Agriculture, as studied in Bustos et al. (2016). That paper shows that this type of technical change leads to a relocation of labor from agriculture to manufacturing, provided that the elasticity of substitution between land and labor (σ) is smaller than the land share in agricultural production. Thus, by combining the insights in Acemoglu (2002) and Bustos et al. (2016) we obtain that, under the condition stated below, skilled-labor augmenting technical change in agriculture leads to the relocation of low-skilled workers away from agriculture.

Proposition 2. *An increase in A_s in agriculture leads to an absolute decrease in the demand for low skilled workers in agriculture if labor and land are strong complements ($\sigma < \varepsilon\Gamma$).*

Proof. See Appendix C. Note that $\Gamma = \left(\frac{(A_T T_a)^{\frac{\sigma-1}{\sigma}}}{(A_L L_a)^{\frac{\sigma-1}{\sigma}} + (A_T T_a)^{\frac{\sigma-1}{\sigma}}} \right)$ is the share of land in agricultural production, and ε is the elasticity of substitution between high- and low-skilled workers. □

Proposition 2 extends the logic of Bustos et al. (2016) to two types of labor, and in doing so we obtain new insights. With only labor and land in agriculture, labor augmenting technical change may lead to a decrease in the demand of labor only if land and labor are sufficiently strong complements. When there are two labor types, the argument is more nuanced. If one of the labor types becomes more productive, then, on the one hand, firms would like to use more of it if it can substitute the other type of labor. On the other hand, however, firms want to use less labor overall if labor and land are strong complements. As a result, when skill-biased-factor-augmenting technologies (A_s) improve, as may be the case in many developing countries when importing technologies from more developed countries, the demand for unskilled labor in agriculture may decrease. With two labor types, as long as $\varepsilon > 1$, strong complementarity ($\sigma < \varepsilon\Gamma$) is a substantially weaker condition than with just one labor type. The reason for that is that part of the adjustment takes place within labor.

2.4 Industrial specialization

From the view point of the manufacturing sector, the release of low-skilled workers from agriculture is akin to an exogenous increase in the relative supply of labor. Heckshen-Ohlin forces imply that this inflow of low-skilled workers into manufacturing expands the industries that use low-skilled labor more intensively. Industrial specialization matters, as we discuss in section 2.5, because the composition of the manufacturing sector determines the long-run growth rate of the economy.

To investigate the effect of skilled-labor augmenting agricultural technical change on industrial specialization, we start by analyzing how it changes the return to the three factors in the economy, namely: land, high- and low-skill labor. To do so, we need to analyze the zero profit conditions in each sector of activity introduced above.

Lemma 1. *If all three sectors are active, the effect of an increase in skilled-biased-factor-augmenting technology in agriculture (A_s) on wages is mediated by the effect of A_s on local knowledge (K_t^h). In particular:*

$$\frac{\partial \ln w_s}{\partial A_s} = \frac{\partial \ln w_u}{\partial A_s} = \frac{\partial \ln K_t^h}{\partial A_s}$$

and the effect of A_s on land prices is given by:

$$\frac{\partial \ln w_T}{\partial A_s} = \frac{\partial \ln K_t^h}{\partial A_s} + \frac{\theta_{S_a}}{A_s \theta_{T_a}}$$

where θ_{S_a} is the cost share of high-skilled workers and θ_{T_a} is the cost share of land in agriculture.

Proof. See Appendix C. □

Lemma 1 says that when all sectors are active the economy is in an "efficiency corrected" factor price equalization set for labor inputs.

Next, we investigate how an increase in skilled-biased-factor-augmenting technology in agriculture leads to particular patterns of industrial specialization. We summarize our results with the following proposition.

Proposition 3. *Skilled-labor augmenting technical change in agriculture (A_s) leads to an expansion of low-skill intensive manufacturing industries, provided that:*

1. *High- and low-skilled workers are imperfect substitutes (i.e. when $\varepsilon > 1$)*

2. *Land and labor are strong complements (i.e. when $\sigma < \varepsilon\Gamma$)*

3. *Agriculture is not much more intensive in low-skilled labor than the low-skill intensive industry.*

Proof. In Appendix C we provide a proof of this proposition assuming that all sectors are active in equilibrium. □

The intuition for this result follows, essentially, from standard Hecksher-Ohlin international trade theory. To fix ideas, let's first consider what would happen in a simple Hecksher-Ohlin world with only two manufacturing industries. An exogenous increase in low-skilled workers expands the low-skilled intensive industry more than proportionately and shrinks the high-intensive industry. This is the only way to guarantee factor market clearing given the excess supply of unskilled labor. Note that given our assumption of a small open economy, prices are fixed. Hence, if output of the high-skilled intensive good does not change and all the extra low-skilled labor enters the low-skill intensive sector, the marginal product of high-skilled labor would be higher in the low-skilled intensive industry. This means that some high-skilled labor would want to leave the high-skilled intensive industry towards the low-skilled intensive one. As a result, the high-skill intensive industry shrinks and all the low-skilled labor released from agriculture plus some high-skilled labor from the high-skill intensive industry enter the low-skilled intensive industry, expanding its size.

In our context we have three sectors (agriculture, low-skilled intensive manufacturing and high-skill intensive manufacturing), instead of two. In this case, unskilled-labor-saving agricultural technological progress frees unskilled labor. From the point of view of the manufacturing sector, this is equivalent to an increase in the supply of unskilled labor, which according to the discussion in the above paragraph generates an expansion in the unskilled industry and a contraction in the skilled industry. Note, however, that if agriculture is very low-skill intensive (much more than the other two sectors), Rybczynski forces would push the "freed labor" back into agriculture. This is why Proposition 3 requires that agriculture is not much more intensive in low-skilled labor than low-skill intensive manufacturing.

2.5 Endogenous growth

As mentioned before, industrial specialization is important in this model because it determines the growth rate of the economy. Hence, our final result relates industrial composition and economic growth. In particular, we have that:

Proposition 4. *When the following conditions hold:*

1. *High- and low-skilled workers are imperfect substitutes (i.e. when $\varepsilon > 1$)*

2. *Land and labor are strong complements (i.e. when $\sigma < \varepsilon\Gamma$)*

3. *Agriculture is not much more intensive in low-skilled labor than the low-skill intensive industry.*

Skilled-labor-augmenting technical change in agriculture (A_s), results in:

1. *Static gains from increased productivity in the agricultural sector.*

2. *Dynamic losses shaped by the decrease in the incentives to invest in new intermediate varieties for the H-industry.*

In particular, the growth rate of consumption is given by:

$$g_C = \frac{\max\{r^l, r^h\} - \rho}{\eta} \tag{11}$$

The change in gross domestic output is given by:

$$\frac{\partial \ln GDP_t}{\partial A_s} = \underbrace{\omega_a \frac{\partial \ln p_a Q_a}{\partial A_s} + \omega_m^{\ell} \frac{\partial \ln p_m^{\ell} Q_{\ell}}{\partial A_s} + \omega_m^h \frac{\partial \ln Q_h}{\partial A_s}}_{\textit{Static gains/losses}} + \underbrace{\mathbb{1}_{\{r^h \geq r^l\}} \frac{\chi}{\eta} \frac{\partial F_m^h}{\partial A_s} t}_{\textit{Dynamic gains/losses}} \tag{12}$$

where $\omega_j = \frac{p_j Q_j}{p_a Q_a + p_m^{\ell} \varsigma Q_m^{\ell} + \varsigma F_m^h}$.

Proof. See Appendix C. □

To provide some intuition for this result note that in equilibrium output growth depends on the sectoral composition of the economy. This is because when agricultural productivity growth pushes workers out of agriculture, the L-industry expands as it absorbs these workers, and its demand for intermediate goods increases. Then, returns for inventing new input varieties used by the L-industry increase and R&D firms direct their innovation efforts towards this industry and K_t^l grows and K_t^h stops growing. However, only the expansion of varieties in the H-industry generates larger productivity and knowledge spillovers. As a result, manufacturing productivity stops growing for a few periods, which depends on how fast labor relocates towards the L industry – something we do not explicitly model given our focus on steady-state outcomes. We labeled the productivity slow down as *dynamic losses*. On impact, however, total output increases since there are productivity gains in agriculture and employment gains in the L-industry, where the set of intermediates also expands. This is what we labeled as *static gains*, which is different from the static gains emphasized in prior literature and that we abstract from in the model.[13]

We provide a qualitative illustration of Proposition 4 in Figure 1. The graph on the left shows the level of profits when inventing for the H- and L-industries. Before the increase in agricultural productivity, profits for inventing in each industry are the same, which we assume in the figure at 5%. When low-skilled workers enter the L-industry it becomes more profitable to invent new varieties for this industry. Hence, for a while intermediate goods producers only invest in expanding the set of intermediates in the L-industry. As the number of intermediates expands, profits decline, up to the point where profits are at the level that they were before the increase in agricultural productivity.

The graph on the right of Figure 1 shows the evolution of the overall output. Shown in a solid line, total output keeps increasing over time (log) linearly at the steady state growth rate. If A_s increases (permanently) at a point in time (denoted by $t = 0$ in the graph), then total output increases instantaneously, as shown by the dashed line. This instantaneous increase is the result of the higher productivity in agriculture (higher A_s) and the increased output in manufacturing due to the entry of low-skilled workers into the sector. However, because the sector that absorbs labor is the L-industry, intermediate goods producers start inventing intermediate varieties for an industry that does not generate productivity growth nor local spillovers. Hence, local productivity stops growing for a few periods.

Note that this model is ambiguous on whether the increased productivity in agriculture is good or bad for long-term output growth. If adjustment is fast, then the amount of time that it takes to bring the profits of inventing for the L-industry down to the steady-state level will be short, and hence the economy may start growing at the steady-state rate at a level that is higher than without technical change in agriculture. Alternatively, it may be that the economy takes a long time to adjust, and hence the level of output is lower than without the technical change in agriculture, as depicted in Figure 1.

Figure 1 goes around here

3 Empirics

In this section we test the predictions of the model using data from Brazil. Our empirical analysis has three main objectives. First, we trace the reallocation of workers with

[13] Previous literature (see Caselli 2005, Restuccia et al. 2008, Lagakos and Waugh 2013, Lagakos and Waugh 2013, or Gollin et al. 2014) argues that there are frictions to mobility from agriculture to manufacturing that impede workers to move across sectors. Instead, in this paper we observe patterns that are in-line with relatively flexible cross-sector mobility, and the static gains come exclusively from increases in agricultural productivity.

different skills across sectors following the introduction of a new labor saving technology in agriculture. Second, we study the implications of this reallocation on local industrial specialization. Third, we document the effect of industrial specialization on innovative activities and productivity growth in the short and medium run.

To establish the direction of causality, from agriculture to manufacturing, we exploit the legalization of genetically engineered (GE) soy in Brazil as a natural experiment. We start by providing background information on GE soy in section 3.1. Our identification strategy uses the potential increase in soy yields that can be obtained with GE seeds in each region based on its weather and soil characteristics as a plausibly exogenous measure of technical change. We describe this strategy in detail, along with the data used to implement it, in sections 3.2 and 3.3. We then develop the three steps of our empirical analysis.

We start by studying the effect of the introduction of GE soy on the reallocation of workers with different skills from agriculture to manufacturing. Our theoretical framework predicts that the adoption of technologies that increase the productivity of skilled labor in the production function – such as GE soy – should displace unskilled workers from agriculture (Proposition 2). We test this prediction using the Population Census, which contains detailed information on both formal and informal workers, and the social security data from RAIS, which contains detailed information on formal employment at yearly frequency. These results are discussed in section 3.4.1.

Next, we study the consequences of this reallocation of unskilled labor from agriculture to manufacturing for the industrial composition of the local economy. Our model predicts that an increase in the relative supply of unskilled labor should be absorbed by industries that use unskilled labor intensively (Proposition 3). These industries correspond to the L-industry in our theoretical framework. We use the Population Census and RAIS data to study the effect of soy technical change on labor allocation across industries within the manufacturing sector, as documented in section 3.4.2.

Finally, we study the impact of industrial specialization on innovation and manufacturing productivity growth. Our theoretical framework predicts that manufacturing productivity should slow down following a large inflow of workers into the L-industry (Proposition 4). In our model, this reallocation makes more profitable for R&D firms to invent new inputs for the L-industry than for the H-industry. However, new input varieties in the L-industry do not increase productivity nor generate knowledge spillovers. As a result, productivity growth declines. In the empirical analysis we measure R&D investments using the wage bill of workers employed in innovative activities. For this purpose, we construct a new measure of labor employed in innovative activities that varies across regions and across sectors using the detailed description of occupations reported in the social security data. We then use this measure to test the mechanism emphasized by the model and study differences in the evolution of manufacturing productivity across regions that were differently exposed to soy technical change in section 3.4.3.

3.1 Background Information on GE Soy

In this section we describe the technological change introduced in Brazilian agriculture by GE soybean seeds. GE soy seeds are genetically engineered in order to resist a specific herbicide (glyphosate). The use of these seeds allows farmers to spray their fields with glyphosate without harming soy plants, reducing labor requirements for weed control.[14] For example, the planting of traditional seeds is preceded by soil preparation in the form of tillage, the operation of removing the weeds in the seedbed that would otherwise crowd out the crop or compete with it for water and nutrients. In contrast, planting GE soy seeds requires no tillage, as the application of herbicide selectively eliminates all unwanted weeds without harming the crop. As activities related to weed control are mostly performed by unskilled workers, the introduction of GE soy seeds tends to displace unskilled labor relatively more than skilled labor.

The first generation of GE soy seeds (Monsanto's Roundup Ready) was commercially released in the U.S. in 1996 and legalized in Brazil in 2003.[15] Prior to 2003, smuggling of GE soy seeds from Argentina was only detected in 2001 and 2002 according to the Foreign Agricultural Service of the United States Department of Agriculture (USDA, 2001). The 2006 Brazilian Agricultural Census reports that, only three years after their legalization, 46.4% of Brazilian farmers producing soy were using GE seeds with the "objective of reducing production costs" (IBGE 2006, p.144). According to the Foreign Agricultural Service of the USDA, by the 2011-2012 harvesting season, GE soy seeds covered 85% of the area planted with soy in Brazil (USDA 2012).

Panel (a) of Figure 2 documents that the legalization of GE soy seeds was followed by a fast expansion of the area planted with soy, which increased from 11 to 19 million hectares between 2000 and 2010.[16] This graph suggests that the area planted with soy started to increase very rapidly already in 2002. Panel (b) of Figure 2 documents that, in the same period, the number of workers employed in the soy sector decreased substantially. This is consistent with the adoption of GE seeds reducing the number of agricultural workers per hectare required to cultivate soy. Bustos et al. (2016) document that labor intensity in soy production fell from 28.6 workers per 1000 hectares in 1996 to 17.1 workers per 1000 hectares in 2006. In addition, the production of soy is less labor-intensive than all other major agricultural activities. According to the Agricultural Census, the average labor intensity of cereals in 2006 was 94.9 workers per 1000 hectares, 129.8 for other seasonal

[14] Other advantages of GE soy seeds are that they require fewer herbicide applications (Duffy and Smith 2001; Fernandez-Cornejo, Klotz-Ingram, and Jans 2002), allow a higher density of the crop on the field (Huggins and Reganold 2008) and reduce the time between cultivation and harvest.

[15] See Law 10.688 of 2003 and Law 11.105 – the New Bio-Safety Law – of 2005 (art. 35).

[16] According to the two most recent agricultural censuses, the area planted with soy increased from 9.2 to 15.6 million hectares between 1996 and 2006 (IBGE 2006, p.144).

crops, and 126.7 for permanent crops.[17] Thus, whenever soy displaced other agricultural activities, labor intensity in agriculture decreased.

Figure 2 goes around here

In Panel (c) of Figure 2, we decompose the decrease in employment in the soy sector between skilled workers and unskilled workers, where a worker is considered as skilled if she has completed at least the 8^{th} grade. As shown, the decrease in employment in the soy sector is entirely driven by low-skilled workers, while the skilled ones were retained. This is consistent with GE soy seeds being an unskilled labor saving technology. Notice that in addition to being less labor intensive, soy production is also more skill intensive than most other agricultural activities. As shown in Panel (d) of Figure 2, the share of skilled workers (those completed at least the 8^{th} grade) employed in soy is above 20 percent, while in most other agricultural activities this share ranges between 5 and 15 percent. Thus, whenever soy displaced other agricultural activities, skill-intensity of agriculture increased.

3.2 Identification strategy

Our identification strategy builds on Bustos et al. (2016): we exploit the legalization of GE soy seeds in Brazil as a source of time variation and differences in the potential increase in soy yields from the introduction of the new technology across regions as a source of cross-sectional variation. The potential increase in soy yields due to GE soy seeds is constructed using data on potential soy yields sourced from the FAO-GAEZ database. This dataset reports the maximum attainable yield for a specific crop in a given geographical area. In addition, it reports the maximum attainable yields of each crop under different technologies or input combinations. Yields under the *low* technology are described as those obtained planting traditional seeds, with no use of chemicals or mechanization. Yields under the *high* technology are obtained using improved high-yielding varieties, with optimum application of fertilizers, herbicides, and mechanization.

Following Bustos et al. (2016), we define technical change in soy production as the difference in potential yields between *high* and *low* technology. This measure aims at capturing the theoretical change in soy yields obtained by switching from traditional soy production to the use of improved seeds and optimum weed control, among other characteristics. Technical change in soy production in micro-region k is therefore defined as:

[17]According to the 2006 Agricultural Census, even cattle ranching uses more workers per unit of land than soy production (30.6 per 1000 hectares).

$$\Delta A_k^{soy} = A_k^{soy,High} - A_k^{soy,Low}$$

where $A_k^{soy,Low}$ is equal to the potential soy yield under the *low* technology and $A_k^{soy,High}$ is equal to the potential soy yield under the *high* technology.[18] ΔA_k^{soy} is our exogenous measure of agricultural technical change in agriculture.

Figure 3 shows the geographical variation in this measure of technical change across micro-regions.

Figure 3 goes around here

The map suggests large variation in agricultural technical change across Brazilian micro-regions. Some regions, most notably the regions around the Amazon river, and near the South-East coast, experienced little changes in soy productivity. Instead, the regions of the Center-West and South gained substantially from the introduction of the new seed.

With decennial data, we use the following specification to estimate the effect of soy technical change on (long-run) changes in outcomes of interest:

$$\Delta Y_k = \alpha + \beta \Delta A_k^{soy} + \varphi X_k + \varepsilon_k \tag{13}$$

where ΔY_k is the change in the outcome of interest in micro-region k between 2000 and 2010 – the years of the last two Population Censuses –, and X_k is a vector of controls of micro-region k. Our identification strategy relies on the fact that the new GE soybeans seeds were introduced around 2001 or 2002 and legalized in Brazil in 2003, and that this new technology disproportionately favored micro-regions with certain soil and weather characteristics (as captured by ΔA_k^{soy}), something that was not anticipated as of 2000. In all our specifications we include the share of rural population in 1991 and the measure of maize technical change presented in Table 1 as baseline controls in order to capture differential trends between urban and rural micro-regions and contemporaneous agricultural changes. In addition, in all our specifications we include macro-region fixed effects, to account for differential trends across the five major geographical regions of the country: north, northeast, south, southeast and central-west. In our extended specification, we also control for the initial level of income per capita, alphabetization rate, and population density at the micro-region level, all observed in 1991 and sourced from the Population Census. These additional controls are meant to flexibly capture differential trends across micro-regions with different initial levels of income and human capital.

[18]Although soy farming in certain areas of Brazil was already using relatively advanced techniques before the introduction of GE soybeans, our conversations with researchers in charge of the FAO-GAEZ dataset show that GE soy seeds are, in fact, the improved seed varieties used to compute predicted soy yields for Brazil under *high* inputs. The predictive power of the instrument on GE soy seeds adoption documented in what follows supports this.

When we analyze the manufacturing sector in detail we use annual data from the social security records (RAIS) and the yearly manufacturing survey (PIA). This allows us to trace the timing of the effect more precisely by estimating two types of equations. First, to provide visual support to our evidence, we estimate the following event-study specification:

$$\ln y_{k,t} = \delta_t + \delta_k + \sum_{j=1999}^{j=2009} \beta_j \Delta A_k^{soy} + \gamma X_{k,t} + t \times X'_{k,1991}\omega + \varepsilon_{k,t} \tag{14}$$

where ΔA_k^{soy} is the long-run change in our exogenous measure of technical change in soy in micro-region k,[19] and $\ln y_{k,t}$ is an outcome of interest in micro-region k at time t. β_j estimates the effect of the change in the productivity of soy in each year between 1999 and 2009. Thus, we flexibly allow β_j to capture the effect of soy technical change on the outcomes of interest in each year. This type of specification is informative of the timing and persistence of the effects. δ_k and δ_t are micro-region and year fixed effects, respectively. $X_{k,t}$ are time-varying controls and $X_{k,1991}$ are the baseline controls discussed above interacted with a time trend. .

With annual data, we estimate the effect of agricultural technical change on manufacturing outcomes using the following specification:

$$\ln y_{k,t} = \delta_t + \delta_k + \beta A_{k,t}^{soy} + \gamma X_{k,t} + t \times X'_{k,1991}\omega + \varepsilon_{k,t}$$

where $A_{k,t}^{soy}$ is defined as potential soy yield under high inputs for the years between 2003 and 2009, and the potential soy yield under low inputs for the years between 1999 and 2002 in micro-region k. δ_k and δ_t are micro-region and year fixed effects, respectively, and $X_{k,t}$ are time-varying controls and $X_{k,1991}$ are baseline controls interacted with a time trend. Hence, β is the (continuous) difference-in-difference estimate obtained from comparing micro-regions before and after 2003.[20]

Table 1 reports a set of results aimed at validating our measure of soy technical change using data from the 1996 and 2006 Agricultural Censuses. First, in Panel A, we show that our measure of soy technical change strongly predicts variation in the actual adoption of GE seeds by Brazilian farmers across micro-regions (columns 1 and 2). Importantly, it does not predict the expansion of area farmed with traditional soy (columns 3 and 4). This indicates that this measure of the effect of technical change on in potential soy yields is a good proxy of the actual benefits of GE soy adoption given soil and weather characteristics of different areas. Second, in Panel B, we show that our measure of soy technical change predicts the expansion of agricultural area farmed with soy, but not the one farmed with maize, the other main temporary crop which experienced significant

[19]The same measure used Equation 13.

[20]In these specifications we use a balanced panel of micro-regions that includes all the micro-regions for which we have observations in each year of the decade.

technological innovation in this period (columns 1 and 2).[21] If we build a measure of maize technical change using the same methodology, we find that such measure predicts the expansion in maize area between 1996 and 2006, but not the expansion of soy area (columns 3 and 4). This indicates that our measure of technical change is a good proxy of technological innovation at the crop level. Note that the results reported in Table 1 effectively replicate the results presented in Bustos et al. (2016) at a larger level of aggregation (micro-region instead of municipality).

Table 1 goes around here

3.3 Data sources

In this section we describe the main data sets used in the empirical analysis. We obtain information on employment from two different sources: the Population Census and RAIS, the social security records dataset of the Ministry of Labor. The Population Census has the advantage of covering both formal and informal workers, and it is available at ten year intervals. RAIS covers only formal employees, but it has the advantage of being available at yearly level. We also use data from two different manufacturing surveys: PIA and PINTEC. We use data from PIA – the Brazilian manufacturing survey – to construct measures of manufacturing productivity. We use data from PINTEC – the Brazilian Innovation survey – to classify industries by innovation intensity. In what follows we describe these four data sources in more detail.

We use the Censuses of 2000 and 2010 to obtain detailed information on employment and wages in all sectors. We focus on individuals with strong labor force attachment. In particular, we include individuals aged between 25 and 55 that work more than 35 hours a week.[22] Differently from social security data, the Population Census covers both formal and informal workers, which makes it well suited to study movements of workers in the agricultural sector – whose labor force is largely informal – as well as any effect on informal employment in manufacturing. For each individual, we define the sector of occupation as the sector of their main job during the reference week of the census. The Population Census also provides information on the number of hours worked during the reference week and the monthly wage.[23] We use information on education to categorize individuals as unskilled or skilled. We define a worker as skilled if they have completed at least the 8^{th} grade, although our results are robust to alternative definitions of this

[21] See Bustos et al. (2016) for a detailed discussion of second-season maize.

[22] In order to deal with extreme observations, we focus on individuals whose absolute and hourly wages are between the 1st and the 99th percentile for the distribution of wages in their respective year, and who work less than the 99th percentile of hours. Moreover, we only consider individuals not enrolled in the education system at the time of the survey.

[23] We compute hourly wages as the monthly wage divided by 4.33 times the hours worked reference week.

threshold. This level should be attained when an individual is 14 or 15 years old and is equivalent to graduating from middle school in the US. We also use data from the Population Census to compute “composition-adjusted” wages (i.e., wages net of observable worker’s characteristics). To this end, we estimate a Mincerian regression of log hourly wages on observable characteristics for the two census years of 2000 and 2010, as explained in Appendix B.

The Annual Social Information System (RAIS) is an employer-employee dataset that provides individual information on the universe of formal workers in Brazil.[24] We use RAIS to study movements of workers across industries within manufacturing at yearly level from 1998 to 2009. As in the Population Census, we focus on individuals aged between 25 and 55 that work more than 35 hours a week.[25] RAIS contains detailed information on workers’ occupations, which we use to construct the new spatial measure of the labor input in innovation activities described below.

We use data from the two manufacturing surveys. We use data on number of workers, value added and wage bill from the Annual Industrial Survey (PIA) to construct our measure of manufacturing productivity.[26] The data from PIA comes aggregated at micro-region and industry level and is constructed using manufacturing firms with more than 30 employees. Since all firms with 30 or more employees are sampled in the PIA survey, our sample is representative at the micro-region and industry level. We focus on firms operating in manufacturing as defined by the CNAE 1.0 classification (codes between 15 and 37) and on the period between 2000 and 2009.

Finally, we use data from the Survey of Innovation PINTEC to classify manufacturing industries by R&D intensity. This survey is designed to capture innovation activities of Brazilian firms and it is available every 3 years starting in 2000. The PINTEC survey provides information on expenditure in R&D at industry level. Using this data we construct a measure of R&D intensity at industry level, measured as the monetary value of R&D expenditures divided by sales in the baseline year 2000. We define *high* R&D intensive industries as those above the median level of R&D intensity, weighting industries by their employment at baseline. Table A.1 reports the full list of manufacturing industries by

[24] Employers are required by law to provide detailed worker information to the Ministry of Labor. See Decree n. 76.900, December 23^{rd} 1975. Failure to report can result in fines. RAIS is used by the Brazilian Ministry of Labor to identify workers entitled to unemployment benefits (*Seguro Desemprego*) and federal wage supplement program (*Abono Salarial*).

[25] Following Helpman, Itskhoki, Muendler, and Redding (2017), our data cleaning procedure includes: (i) restricting to workers employed as of December 31st in each year; (ii) restricting to the highest-paying job for each worker that appears more than once in the data during one year (randomly dropping ties).

[26] We define employment as end-of-year number of workers, and value added as the difference between output value and production costs. Specifically, the value of output is defined as the sum of revenue from industrial sales, the value of production used for investment and the changes in inventories, whereas production costs are equal to the sum of the cost of industrial operations and the cost of materials used.

R&D intensity and skill intensity. R&D intensity and skill intensity at industry level are highly correlated, as can be seen in Figure A.1 in the Appendix.[27]

Table 2 reports summary statistics of individual level characteristics observed in the Population Census for workers operating in agriculture, low-R&D manufacturing, high-R&D manufacturing and services.[28] As shown, there is large heterogeneity in skill intensity of workers across these broad sectors. Almost 90 percent of workers in agriculture had not completed the 8^{th} grade in 2000, while this number is around 50 percent for manufacturing and services. Within manufacturing there are also large differences, where the share of high-skill workers tend to be higher in high R&D industries, particularly in 2010.

Table 2 goes around here

Table 3 provides summary statistics for the main variables used in the empirical analysis at the micro-region level. micro-regions are statistical units defined by the Brazilian Statistical Institute (IBGE) and consist of a group of municipalities. There are 557 micro-regions in Brazil, with an average population of around 300,000 inhabitants. We use micro-regions as an approximation of the local labor market of a Brazilian worker. They can be thought of as small, open economies that trade in agricultural and manufacturing goods but where production factors are immobile.[29] For outcomes sourced from the Population Census – which are observed in 2000 and 2010 — we report the mean and standard deviation of their level in the baseline year (2000) and of their change between 2000 and 2010.

Table 3 goes around here

A new measure of innovation across space

In the model, the increase in the relative size of low-skill industries reduces incentives to innovate in high-skill industries, the only ones which generate productivity growth and knowledge spillovers, thus aggregate growth slows down. To test this prediction we need to

[27]Notice that data on R&D expenditure from the PINTEC survey is not representative at the micro-region level. Thus, to construct a measure of innovation that is representative at any geographical level, we use the description of occupations reported in the social security records, as described below.

[28]We define agriculture, manufacturing and services by following the classification of the CNAE Domiciliar of the 2000 census. Agriculture includes Sections A and B (agriculture, cattle, forestry, and fishing). Manufacturing includes Section D, which corresponds to the transformation industries. Services include: construction, commerce, lodging and restaurants, transport, finance, housing services, domestic workers, and other personal services. We exclude the following sectors because they are mostly under government control: public administration, education, health, international organizations, extraction, and public utilities.

[29]In Table A.3 of the Appendix we show that internal migration did not respond to the shock. This is in line with evidence from Brazil's lack of internal migration responses documented also in Dix-Carneiro and Kovak (2019) and Costa, Garred, and Pessoa (2016).

observe innovation at the micro-region level, our unit of observation for empirical analysis. For this purpose, we develop a new measure of innovation which is representative at any level of geographical disaggregation using the description of occupations in social security data (RAIS). More specifically, we propose a new measure of the labor input in innovation activities based on textual analysis of the task descriptions of more than 2500 occupations. Tasks generating innovations include, for example, developing new products and processes, creating prototypes, or optimizing methods of production. An important advantage of this innovation measure is that it allows us to track innovation workers across sectors and regions. This is because the social security data covers the universe of formal firms. In contrast, standard manufacturing innovation surveys, such as PINTEC, are based on a sample of firms which is not representative at low levels of geographical disaggregation. An alternative measure of innovation that is potentially representative at micro-region level is patents. However, as we detail below, patenting is relatively uncommon and is only a small part of innovation activities in Brazil. In addition, patented innovations are more likely to generate spillovers beyond the locality where R&D activities take place, while the focus of our empirical analysis are local productivity spillovers.

In what follows we describe our methodology to identify workers in innovative occupations within each manufacturing industry and region. As a first step, we digitized the text containing the official description of the tasks associated with each occupation as provided by the Ministry of Labor. In the second step, we defined a set of 39 keywords or combination of keywords capturing tasks related to innovative activities. To generate this list, we relied on keywords – or combination of keywords – that are used to define innovative activities in the technical documentation of PINTEC, the Survey of Innovation of Brazilian firms. The list of keywords is reported in Appendix Table A.2. As shown, most entries are a combination of a verb and a noun describing a task associated with innovation. These combinations can be grouped in those capturing innovation of products (e.g. "develop/improve product/s"), innovation of processes (e.g. "develop/improve/test process/es"), innovation of machinery and equipment (e.g. "develop device/s, "develop equipment"). We also include single nouns, combinations of nouns or combinations of nouns and adjectives that are often found in the description of innovation intensive tasks (e.g. "innovation", "prototypes", "research and development", "new technologies"). Finally, in the last step, we run a text analysis that identifies all occupations whose description contains at least one of the keywords listed in Appendix Table A.2. This methodology identifies 251 occupations, which we define as innovation-intensive.

Figure A.2 shows the total number and the share of manufacturing workers in innovation-intensive occupations in Brazil. According to our measure, the number of workers in innovation-intensive occupations increased from approximately one hundred thousand in 2000 to three hundred thousands in 2014, and started falling afterward when Brazil entered into a severe recession. Workers in innovation intensive occupations constitute between 3 and 4 percent of total manufacturing formal employment. This share has been increasing during the period under study from 2.5 percent in the early 2000s to slightly above 4 percent in most recent years.[30]

In Figure A.3 of the Appendix we report the correlation between our measure of employment share in innovative occupations at the industry level and other sector level measures of innovation. As shown, our measure has a positive correlation with other input based measures of innovation, such as R&D expenditure per worker from PINTEC. Our measure has the advantage of capturing innovation effort that is often not categorized under "R&D", especially by smaller firms. More importantly, as mentioned above, our measure allows to capture innovation effort not only at the sector level but also at fine geographical level. Figure A.4 reports the share of local manufacturing employment engaged in innovation intensive activities in each micro-region of Brazil in the baseline year 2000. As shown, the share of innovation workers ranges from 0 to almost 20 percent of formal manufacturing employment, with higher shares observed in the coastal regions of the south and south-east of Brazil, but also in several micro-regions encompassing large cities in the north and western regions of the country.

Finally, we want to briefly discuss the differences between our measure of innovation based on workers' task description and other measures of innovation based on the outcome of the innovation process, such as patenting. The advantage of using patent data to measure innovation is that – differently from input based measures – patents capture the efficiency of the innovation process and allow researchers to make statements about the quality of the innovation produced, e.g. by using patent citations (Carlino and Kerr 2015). However, one important disadvantage of patent data is that many innovations are not patented. Data from PINTEC, the survey of innovation of Brazilian firms, shows that, in the decade 1997 to 2008, 34 percent of surveyed firms declare to have introduced some innovation, such as new processes or products. However, only 7 percent of those firms have filed a patent application or have an approved patent for such innovation.[31] Thus, one advantage of our measure relative to data on patenting activities is that patenting is relatively uncommon and is only a small part of innovation activities. This is not only the case in Brazil as survey data from approximately 1500 R&D labs of manufacturing firms in the US, shows that patenting is used less frequently than other approaches to protect the return from invention, as patent applications require firms to disclose to competitors

[30]The Brazilian Ministry of Labor has updated its classification of occupations in 2002. RAIS uses the new classification (CBO2002) starting from 2003. We identify innovation intensive occupations using the the description of tasks provided for the CBO2002 classification. To extend our analysis to the pre-2003 years we match the old classification (CBO 1994) and new classification (CBO 2002) using the official correspondences provided by the Ministry of Labor. Whenever one occupation in the old classification is matched with multiple occupations in the new one, we weight the number of workers in that occupation by the share of innovation workers observed in the first year in which the new classification is used (2003).

[31]These statistics are based on Table 6497 of the PINTEC surveys run in 2000, 2003, 2005 and 2008. Each PINTEC survey captures the innovative activities in the previous three years, so they effectively cover the decade 1997 to 2008. The statistics reported are averages across the four waves.

a large amount of information Cohen, Nelson, and Walsh (2000). According to the same survey, smaller firms tend not to apply for patents due to their legal costs, and are also more likely to consider patents ineffective.

3.4 Estimates of the effects of agricultural technical change

In this section we estimate the effects of agricultural technical change on the allocation of workers with different skills across the agricultural, manufacturing and service sectors. Next, we test the predictions of the model regarding the effects of the reallocation of unskilled labor towards manufacturing on industrial specialization, innovation and productivity growth in the short and medium run.

3.4.1 Labor reallocation across sectors

We start in Table 4 using Census data to document that soy technical change generated a reallocation of (mostly unskilled) labor from away from agriculture into manufacturing. Panel A shows that micro-regions with higher exposure to soy technical change experienced a decrease in the share of workers employed in agriculture and an increase in the share of workers employed in manufacturing and services.[32] The estimate presented in column (2) indicates that micro-regions with a one standard deviation larger increase in soy technical change experienced a 2.4 percentage points larger decline in agricultural employment share. This estimate is stable to the inclusion of initial controls described in section 3.2. Agricultural workers displaced by the new technology relocated mostly into manufacturing. Manufacturing employment share increased by 1.8 percentage points for a standard deviation differential change in soy technical change, while services employment share increased by 0.6 percentage points. Hence, the results presented in Panel A, indicate that soy technical change was labor-saving and led to structural transformation, which are the main findings documented in Bustos et al. (2016).[33]

Table 4 goes around here

[32]Soy technical change had only small and not significant effects on total employment. Thus, the employment changes that we document in what follows are not driven by migration between micro-regions or by changes in the total number of workers employed, but by movement of workers across sectors within micro-regions. In Table A.3 in the Appendix we provide evidence on the effect of soy technical change on total employment and migration.

[33]Bustos et al. (2016) find that soy technical change had a positive and significant effect on the employment share in manufacturing but no significant effect on the employment share in the services sector. Panel A of Table 4 in this paper documents that micro-regions more exposed to soy technical change experienced an increase in employment share in both manufacturing and services. There are two reasons behind this difference in results when the outcome is the employment share in the services sector. The first is that, in this paper, we focus on remunerated labor – i.e. workers receiving a wage – whereas Bustos et al. (2016) also included workers who helped household members without receiving a payment or worked in subsistence agriculture. The second is the unit of observation, which is a micro-region in this paper, a municipality in Bustos et al. (2016).

In Panels B and C of Table 4, we study the effect of soy technical change on the reallocation of workers with different skills across sectors. More specifically, we characterize whether the reallocation of workers from agriculture to manufacturing documented in Table 4 is mostly driven by unskilled or skilled workers. In Panel B of Table 4 we focus on unskilled workers. We find that micro-regions more exposed to soy technical change experienced a reallocation of unskilled workers from agriculture to manufacturing. The magnitude of the estimated coefficients indicate that micro-regions with a standard deviation higher increase in soy technical change experienced a 2.4 percentage points larger decrease in the share of low-skilled workers employed in agriculture, and a corresponding 2.2 percentage points larger increase in the share of low-skilled workers employed in manufacturing. These magnitudes correspond to a 7.2 percent decrease in the initial share of low-skilled workers employed in agriculture, and a 16.1 percent increase of the share of those employed in manufacturing. Combined with the fact that soy technical change had no differential effect on total employment (Table A.3 in the Appendix), these results are consistent with a decline in the absolute demand for low-skilled labor in agriculture in response to skilled labor-augmenting technical change, as predicted by the model. Finally, in Panel C we focus on skilled workers. We find that micro-regions more exposed to soy technical change experienced a larger decrease in the share of high-skill workers in agriculture, and a larger increase in the share of high-skill workers employed in manufacturing. In terms of magnitude, the effect of soy technical change on low-skill labor is about twice as large as the effect on high-skill labor.

We explore the labor reallocation process described above in more detail using yearly data from RAIS in what follows. Although RAIS data captures only formal employment, its annual frequency allows us to check whether the employment changes documented with Census data occurred right after GE soy was introduced in Brazil. For this, we plot the interaction of year dummies with our measure of soy technical change as explained in Section 3.2. As can be seen in Figure 4 (a), low-skilled labor started to move towards manufacturing in micro-regions more exposed to soy technical change around 2002, while there is no systematic difference in the trends leading to this year. When focusing on formal employment captured by social security data, the differential increase in labor moving towards manufacturing is almost exclusively driven by unskilled labor, as shown in Figure 4 (b). The timing of the effect suggests that changes were permanent. Reallocation of unskilled labor towards manufacturing started around 2002, one year after the first reported smuggling of the new soy seeds and the year when area planted with soy started expanding at a faster rate (Figure 2). The reallocation then accentuated around 2004, one year after the formal legalization of GE soy in Brazil, and stabilized during the second half of the decade.

Figure 4 goes around here

Taken together, the estimates presented in Table 4 and Figure 4 show that the agricultural sector experienced a decrease in its employment share of both low-skill and high-skill labor, while the manufacturing sector experienced an increase in employment driven mainly by low-skill labor. These findings indicate that labor-saving technical change in agriculture driven by the adoption of GE soy was skill-biased and led mainly low-skill workers to reallocate towards manufacturing.[34]

3.4.2 Industrial specialization

As discussed in Section 2, our model predicts that if the supply of low-skilled workers released from agriculture is absorbed into manufacturing, this sector tends to specialize in unskilled labor-intensive industries. In the model, these industries are also the least innovative in the sense that the introduction of new intermediate inputs do not generate productivity nor knowledge spillovers. This assumption is consistent with the data where we observe that unskilled-labor intensive industries are also the least innovative as measured by R&D expenditure per unit of output, as discussed in 3.3. In what follows, we study industrial specialization by splitting manufacturing industries between low and high-R&D as described in Section 3.3. We obtain similar results when splitting industries by skill intensity, as shown in the Appendix.

We start by investigating the effect of soy technical change on industrial specialization using Census data. The results are reported in Table 5. In Panel A we decompose the effect of soy technical change on the manufacturing employment share between high- and low-R&D intensive industries. Column (1) replicates the estimate shown in Table 4. Columns (2) and (3) show that 84% (.21 / .25) of the soy-driven increase in the share of workers in manufacturing occurs in industries with low levels of R&D intensity. In Panels B and C, we report the estimates separately for high and low-skilled workers. The numbers indicate that almost all workers moving into manufacturing – both unskilled and skilled –found employment in the manufacturing industries with lower R&D intensity.

Table 5 goes around here

Our model predicts that low-skill workers reallocate to sectors whose skill intensity is sufficiently close to that of agriculture. To investigate this further, we split manufacturing industries into four quartiles of R&D intensity, each employing one fourth of total

[34] In the Appendix A we provide additional evidence that supports our findings. In particular, Tables A.4 and A.5 show that wages of high skill workers increased while those low-skill workers did not change, consistent with an increase in the relative demand for high-skill labor. We show that the reason why low-skilled wages did not decrease despite the excess supply of workers is related to the increase in minimum wages in Brazil during the decade. The share of workers at the minimum wage increased disproportionately in high soy shocks micro-regions, as documented in Table A.6. This evidence points to the fact that adjustment is easier observed in quantities rather than in prices, hence our focus on labor reallocation throughout the paper.

manufacturing workers at baseline. Then, we estimate which of the four groups absorbed low-skilled labor using the following equation:

$$\Delta \frac{L_{m,ik}}{L_k} = \alpha + \beta_i \Delta A_k^{soy} \times \gamma_i + \gamma_i + \varepsilon_{ik} \quad (15)$$

where i indexes quartiles of R&D intensity at industry level and k indexes micro-regions. The outcome variable in this regression is the change in manufacturing employment in each quartile of industry R&D intensity as a share of total employment in a given micro-region. For example, $\Delta \frac{L_{m,1k}}{L_k}$ is the change in manufacturing workers employed in industries belonging to the lowest quartile of initial R&D-intensity divided by total workers in a given micro-region. When estimating equation (15) we include the standard set of controls at micro-region level interacted with quartiles of R&D intensity at the industry level (γ_i). Figure 5 shows the results, where we report the estimated coefficients on soy technical change by quartile of industry R&D-intensity. The Figure shows that the effect of soy technical change on the change in manufacturing employment share documented in Table 4 is concentrated in industries in the lowest quartile of R&D-intensity. We obtain similar results when splitting industries by skill intensity, as shown in Appendix Figure A.5. Hence, as the model predicts, labor released from agriculture finds employment in manufacturing sectors that are similar to agriculture in terms of skill-intensity.

Figure 5 goes around here

3.4.3 Innovation and manufacturing productivity

In section 3.4.2 we showed that agricultural technical change led to a reallocation of low-skilled workers into low-skill intensive manufacturing industries. A key implication of the theoretical framework presented in Section 2 is that specialization in low-skill intensive industries slows down manufacturing productivity growth. This is because as low-skill-intensive industries expand, the return to introduce new intermediate inputs in these industries increases relative to the high-skill industries. Then, investment in product development reallocates to low-skill industries. However, these are traditional industries with lower scope for generating productivity enhancing innovations. In addition, they do not generate knowledge spillovers towards other industries. As a result, productivity growth slows down both in low and high skill industries.

We can test this specific mechanism thanks to the richness of the social security data which permits to construct a granular measure of investment in innovation. As described in Section 3.3, we use the description of workers' occupations to develop a new measure of employment in innovation-intensive activities that varies both across regions and sectors. This measure allows us to investigate the effect of agricultural technical change on the allocation of innovative activities across industries. As a proxy for the investment in innovative activities in each industry we use the total wage bill of workers in innovation-intensive occupations.[35]

First, we confirm the findings reported above with Census data. Figure 6 reports the results of estimating equation (14) when the outcome variable is the total wage bill of workers employed in innovation and non-innovation intensive occupations (in logs). We start by focusing on workers in non-innovation intensive occupations. Figure 6 (a) shows that regions more exposed to soy technical change experienced an increase in the wage bill in non-innovative activities within low-R&D industries. The timing of the effect is consistent with the timing of the legalization of the new GE soy seeds and with the reallocation of unskilled labor towards low-R&D manufacturing industries documented using Population Census data in the previous section. In turn we do not find any effect of agricultural technical on the wage bill of non-innovative occupations in high R&D industries, which is also consistent with the findings reported above.

Next, we study the effect of agricultural technical change on investment in innovation-intensive activities. We define workers engaged in innovation intensive occupations as those effectively producing new ideas – such as new products and processes – within each industry. As shown in Figure 6 (d), regions more exposed to soy technical change experienced a significant decline in investment in innovative activities within high R&D intensive industries, whose timing corresponds with the legalization of GE soy and the expansion of soy that followed. As shown in panel (c) we find a modest, positive, and non statistically significant effect of soy technical change on innovative activities in low R&D industries. The model provides an interpretation for these findings: the expansion of the low-R&D intensive industries generated higher returns for solving problems or introducing new intermediate inputs in that industry, drawing innovation workers from high-R&D industries. However, as these innovative workers relocate from high to low R&D intensive industries they tend to specialize in tasks that do not generate productivity enhancing innovation nor knowledge spillovers. For example, they might take managerial occupations which are not included in our measure of innovative labor.

Figure 6 goes around here

Table 6 quantifies the effects documented in Figure 6. The coefficient reported in column (1) indicates that micro-regions with a one standard deviation larger increase in potential soy yields experienced an 11 percent higher increase in the wage bill of non-innovative labor in low R&D manufacturing industries. The coefficient in column (2) is

[35]We think of wage bill as a measure of investment in innovative activities in a given sector that captures not just the number of workers employed but also their "quality" (as captured by their remuneration). Our results are robust to using employment instead of wage bill as a measure of investment in innovative activities in each sector.

instead close to zero and not statistically significant, consistent with this inflow of non innovative labor coming from the agricultural sector rather than from other industries within manufacturing. Column (4) shows a negative and significant effect of soy technical change on investment in innovative labor in high R&D intensive industries, as measured by the wage bill. The magnitude of the coefficient indicates that micro-regions with one standard deviation larger increase in soy technical change experienced a 20 percent larger decline in the wage bill of innovative labor in high R&D industries. The coefficient in column (3) is instead positive but not statistically significant. As mentioned above, innovative workers moving to low R&D industries tend to switch to non-innovative activities.

Taken together, the results in Figure 6 and Table 6 are consistent with a reallocation of workers engaged in innovative activities in high R&D to low R&D industries. In our model, this reallocation is driven by the initial movement of unskilled workers from agriculture towards low R&D industries, which increases the demand for investment in new intermediate products those industries. As these new intermediates do not generate productivity improving innovations nor knowledge spillovers, local manufacturing productivity growth slows down. We turn to investigate this prediction next.

Table 6 goes around here

We use equation (14) to estimate the effect of agricultural technical change on labor productivity in manufacturing. We measure labor productivity as value added over wage bill as observed in the Annual Industrial Survey PIA, where we observe these variables aggregated at the micro-region and industry level. Figure 7 reports the results. The graph shows that micro-regions more exposed to soy technical change experienced a relative decline in manufacturing labor productivity. The effect becomes statistically significant and increases in magnitude in the years after the legalization of GE soy in Brazil. In other words, the effect takes place with some delay relative to the effects documented for labor reallocation and investment in innovative activities, as one would perhaps expect given that those results capture the impact of soy technical change on the inputs of the innovation process.

Figure 7 goes around here

While Figure 7 is in line with the predictions of the model, it could also be explained by labor productivity decreasing in manufacturing purely as a result of a composition effect. For example, if labor productivity is lower in low-R&D intensive industries, a movement of workers towards these industries necessarily results in lower aggregate labor productivity. Our model highlights instead that manufacturing productivity decreases because investment in innovation in high-R&D intensive sectors decreases, affecting productivity in both sectors. Thus, in Table 7, we split the effect of soy technical change on labor productivity between high and low R&D intensive industries. As shown in columns (2) and (3), the decrease in manufacturing productivity occurs in both high- and low-R&D intensive industries.

Next, in column (4) of Table 7, we study the effect of soy technical change on the share of innovation performed in high R&D intensive sectors. As shown, we find a negative and significant coefficient on soy technical change, indicating a reallocation of innovative activities out of high R&D intensive sectors. In terms of magnitude, the coefficients reported in columns (1) and (4) of Table 7 indicate that micro-regions with a one standard deviation larger increase in potential soy yields experienced a 3.2 percentage points decline in the share of innovation performed in high R&D intensive industries, and a corresponding 1.4 percent lower yearly growth rate of manufacturing productivity in the period after the legalization of GE soy.[36] As shown in Panel B, these results are similar in magnitude independently of whether we measure labor productivity as value added divided by total wage bill or value added divided by number of workers.

Table 7 goes around here

4 Conclusions

The reallocation of labor from agriculture into manufacturing is generally regarded as positive in economic development literature. Several studies have documented that the manufacturing sector has, on average, higher productivity and pays higher wages. However, little is known about which type of workers are released from the agricultural sector and which manufacturing industries absorb them during the process of structural transformation.

Our paper contributes to the literature by showing that the forces driving structural transformation can shape the type of industries in which a country specializes. In most countries, the process of industrialization can be ascribed to one of two forces: "push" forces, such as new agricultural technologies that push workers out of agriculture, or "pull" forces, such as industrial growth that pull workers into manufacturing. We show that when labor reallocation from agriculture to manufacturing is driven by labor-saving agricultural productivity growth – rather than manufacturing labor demand – it can

[36] The effect on the yearly growth rate of manufacturing productivity is computed by multiplying the coefficient in column (4) of Table 7 by a standard deviation in soy technical change and computing the annualized effect on labor productivity for the post GE soy legalization years.

generate an expansion in low R&D-intensive manufacturing sectors which can reduce investment in innovation and slow down aggregate manufacturing productivity growth.

We guide our empirical analysis through the lenses of an open economy, three sector endogenous growth model. The model suggests that the low-skilled labor released from agriculture should find accommodation in the low-skilled intensive manufacturing industries. The expansion of these industries increases demand for new intermediate inputs which attracts innovation workers from high R&D industries. Once employed by the low-R&D industry, these workers engage in tasks that do not lead to productivity enhancing innovations nor knowledge spillovers. We use yearly data on labor productivity and a new measure of investment in innovation based on task descriptions to show that the data supports the predictions of the model.

Taken together, our findings indicate that structural transformation obtained through labor-saving and skill-biased technical change in agriculture – which may be quite common when developing countries adopt agricultural technologies from more developed ones – can attenuate the standard gains from reallocation into manufacturing emphasized by the existing literature.

References

Acemoglu, D. (2002, October). Directed Technical Change. *The Review of Economic Studies 69*(4), 781–809.

Acemoglu, D. and V. Guerrieri (2008, June). Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy 116*(3), 467–498.

Baumol, W. (1967). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *American Economic Review 57(3)*, 415–26.

Boppart, T. (2014). Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences. *Econometrica*.

Bragança, A. (2014). *Three Essays on Rural Development in Brazil*. Ph. D. thesis, PUC-Rio.

Buera, F., J. Kaboski, and R. Rogerson (2015). Skill-Biased Structural Change. *Mimeo*.

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

Carlino, G. and W. R. Kerr (2015). Agglomeration and innovation. In *Handbook of regional and urban economics*, Volume 5, pp. 349–404. Elsevier.

Caselli, F. (2005). Chapter 9 Accounting for Cross-Country Income Differences. Volume 1, Part A of *Handbook of Economic Growth*, pp. 679 – 741. Elsevier.

Caselli, F. and W. J. Coleman (2001, June). The U.S. Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy 109*(3), 584–616.

Cohen, W. M., R. R. Nelson, and J. P. Walsh (2000). Protecting their intellectual assets: Appropriability conditions and why us manufacturing firms patent (or not).

Costa, F., J. Garred, and J. P. Pessoa (2016, September). Winners and losers from a commodities-for-manufactures trade boom. *Journal of International Economics 102*, 50–69.

Dix-Carneiro, R. and B. K. Kovak (2019, March). Margins of labor market adjustment to trade. *Journal of International Economics 117*, 125–142.

Duffy, M. and D. Smith (2001). "Estimated Costs of Crop Production in Iowa". *Iowa State University Extension Service FM1712*.

Ethier, W. (1982). National and international returns to scale in the modern theory of international trade. *American Economic Review*.

Fernandez-Cornejo, J., C. Klotz-Ingram, and S. Jans (2002). "Estimating Farm-Level Effects of Adopting Herbicide-Tolerant Soybeans in the USA". *Journal of Agricultural and Applied Economics 34*, 149–163.

Ferreira, F., S. Firpo, and J. Messina (2017). Ageing Poorly? Accounting for the Decline in Earnings Inequality in Brazil, 1995-2012. *IDB Working Paper Series*.

Foster, A. D. and M. R. Rosenzweig (1996). Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *The American Economic Review 86*(4), 931–953.

Franck, R. and O. Galor (2019). Flowers of evil? industrialization and long run development. *Journal of Monetary Economics*.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics*.

Gollin, D., S. Parente, and R. Rogerson (2002). The Role of Agriculture in Development. *American Economic Review*, 160–64.

Grossman, G. and E. Helpman (1991). *Innovation and Growth in the Global Economy.* The MIT Press.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2017). Trade and inequality: From theory to estimation. *The Review of Economic Studies 84*(1), 357–405.

Herrendorf, B. and T. Schoellman (2018). Wages, Human Capital and Structural Transformation. *American Economic Journal: Macroeconomics*.

Huggins, D. R. and J. P. Reganold (2008). "No-Till: the Quiet Revolution". *Scientific American 299*, 70–77.

IBGE (2006). *"Censo Agropecuário 2006"*. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística (IBGE).

Imbert, C., M. Seror, Y. Zhang, and Y. Zylberberg (2019). Migrants and firms: Evidence from china.

Kongsamut, P., S. Rebelo, and D. Xie (2001). Beyond Balanced Growth. *The Review of Economic Studies 68*(4), 869–882.

Krugman, P. (1987). The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies. *Journal of Development Economics 27*(1-2), 41–55.

Kuznets, S. (1973). Modern Economic Growth: Findings and Reflections. *American Economic Review*.

Lagakos, D. and M. E. Waugh (2013). Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review 103*(2), 948–80.

Lewis, A. W. (1954). Economic Development with Unlimited Supplies of Labor. *The Manchester School*.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*.

Matsuyama, K. (1992a). A Simple Model of Sectoral Adjustment. *The Review of Economic Studies*, 375–388.

Matsuyama, K. (1992b, December). Agricultural Productivity, Comparative Advantage, and Economic Growth. *Journal of Economic Theory 58*(2), 317–334.

McMillan, M. and D. Rodrik (2011). Globalization, Structural Change and Productivity Growth. In *Making Globalization Socially Sustainable*. ILO/WTO.

McMillan, M., D. Rodrik, and C. Sepulveda (2017). Structural Change, Fundamentals and Growth: A Framework and Case Studies. *National Bureau of Economic Research Working Paper Series No. 23378*.

Murphy, K. M., A. Shleifer, and R. Vishny (1989). Income Distribution, Market Size, and Industrialization. *The Quarterly Journal of Economics 104*(3), 537–564.

Ngai, L. R. and C. A. Pissarides (2007). Structural change in a multisector model of growth. *American Economic Review 97*(1), 429–443.

Nurkse, R. (1953). *Problems of Capital Formation in Underdeveloped Countries*. Oxford: Basil Blackwell.

Peters, M. (2019). Market size and spatial growth - evidence from germany's post-war population expulsions. *Unpublished manuscript*.

Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis. *Journal of Monetary Economics 55*(2), 234 – 250.

Romer, P. (1987). Growth based on increasing returns due to specialization. *American Economic Review*.

Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy 98*(5).

Rostow, W. (1960). *The Stages of Economic Growth: A Non Communist Manifesto*. Cambridge University Press.

Schultz, T. (1953). *The Economic Organization of Agriculture*. New York: McGraw-Hill.

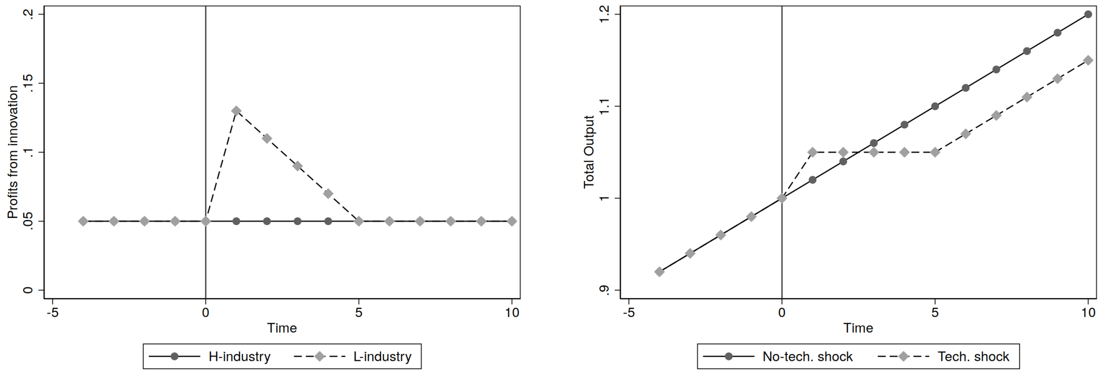
Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. McMaster University Archive for the History of Economic Thought.

USDA (2001). "Agriculture in Brazil and Argentina: Developments and Prospects for Major Field Crops". United States Department of Agriculture, Economic Research Service.

USDA (2012). "Agricultural Biotechnology Annual". United States Department of Agriculture, Economic Research Service.

5 Figures and Tables

Figure 1: Theoretical model: Evolution of intermediate producers' profits and output given an increase in A_s



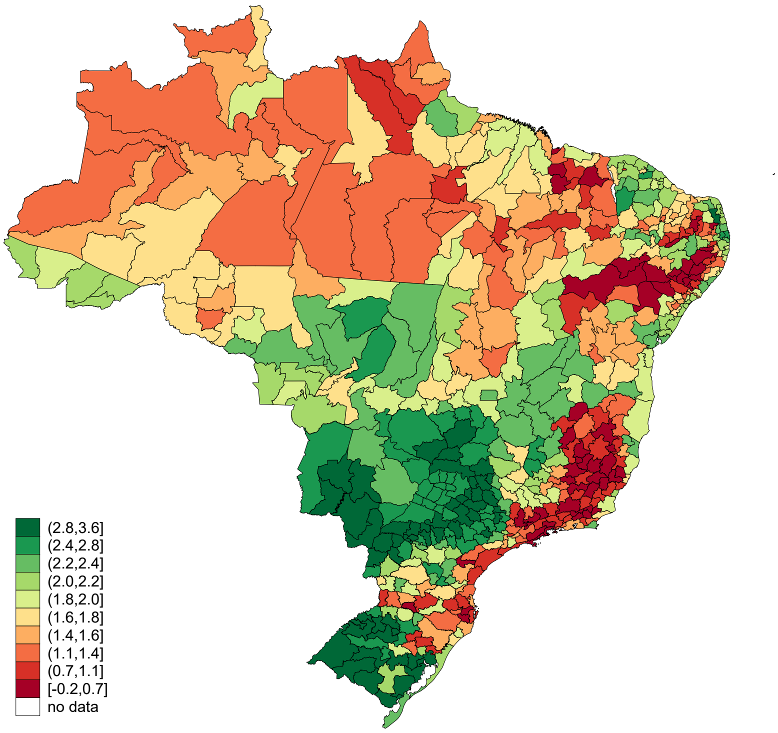
Profits intermediate producers **Total Output**

Notes: This figure shows the qualitative theoretical evolution of the profits of intermediate producers (left panel) and total output (right panel) implied by our model when at time $t = 0$ skilled-biased-factor-augmenting technology (A_s) in agriculture increases. The figure displays the evolution of the economy both with (dashed line) and without (solid line) the technological change.

Figure 2: Soy Production and Employment

Notes: Figures in Panels (a) and (b) are from Bustos et al. (2016). Data sources are CONAB (Panel A), PNAD (Panel B and C) and 2000 Population Census (Panel D). CONAB is the Companhia Nacional de Abastecimento, an agency within the Brazilian Ministry of Agriculture, which runs surveys of farmers and agronomists to monitor the annual harvests of major crops in Brazil. PNAD is the Brazilian National Household Sample Survey. The states of Rondonia, Acre, Amazonas, Roraima, Pará, Amapá, Tocantins, Mato Grosso do Sul, Goias, and Distrito Federal are excluded due to incomplete coverage by PNAD in the early years of the sample. In Panels C and D, an individual is classified as skilled if she has completed at least the 8^{th} grade.

Figure 3: Δ in Potential Soy Yield 2000-2010

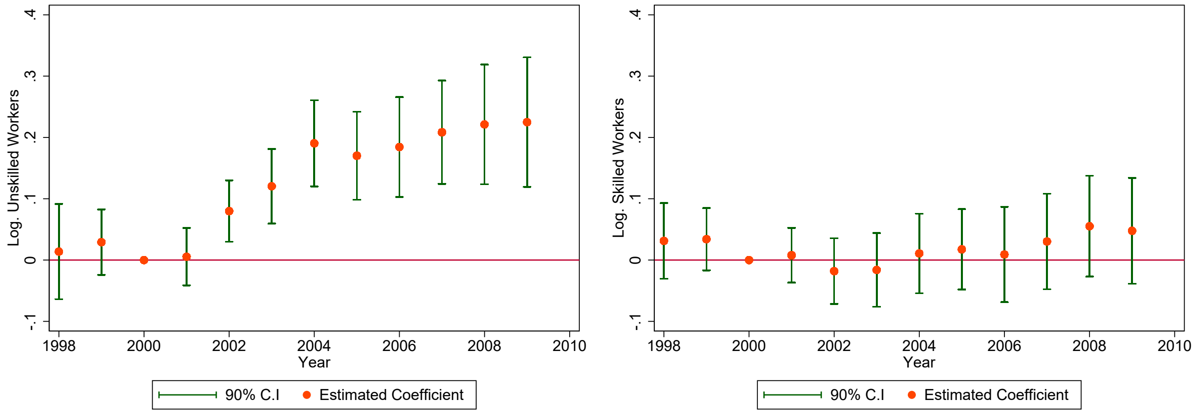


Notes: Authors' calculations from FAO-GAEZ data. Technical change in soy production for each microregion is computed by deducting the average potential yield under low inputs from the average potential yield under high inputs.

Figure 4: Effect of agricultural technical change on manufacturing employment

Yearly Social Security Data (1998-2009)

(a) **Unskilled Labor** (b) **Skilled Labor**



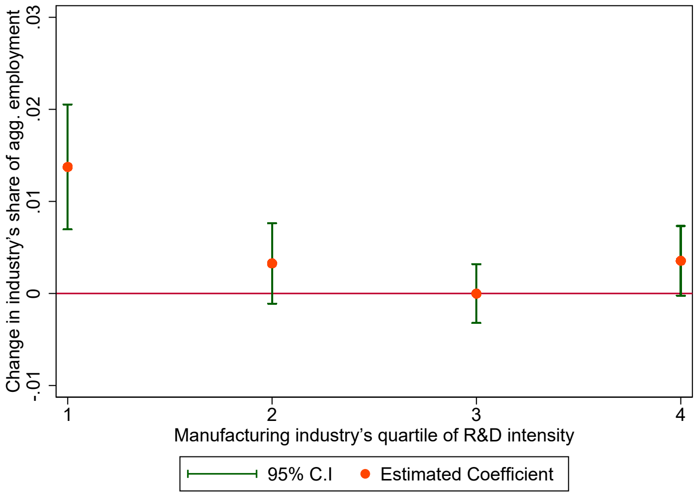
Notes: The plot shows the point estimates and the 90% confidence intervals for the estimates of the β_j coefficients of the following regression:

$$\ln y_{k,r,t} = \delta_t + \delta_k + \delta_{rt} + \sum_{j=1998}^{j=2009} \beta_j \Delta A_k^{soy} + tX'_{k,1991}\omega + \varepsilon_{k,t}$$

Standard errors are clustered at the microregion level. $\ln y_{k,r,t}$ corresponds to aggregate log. employment of skilled and unskilled labor in microregion k located in region r at the end of year t for manufacturing industries. An individual is classified as skilled if she has completed at least the 8^{th} grade. (Source: RAIS).

Figure 5: Effect of agricultural technical change on industrial specialization within manufacturing

Decadal Population Census Data (2000-2010)



Notes: The plot shows the β_i coefficients of the following regression:

$$\Delta \frac{L^k_{m,i}}{L^k} = \alpha + \beta_i \Delta A_{soy} \times \gamma_i + \gamma_i + \varphi X_{k,1991} + \varepsilon^i_k$$

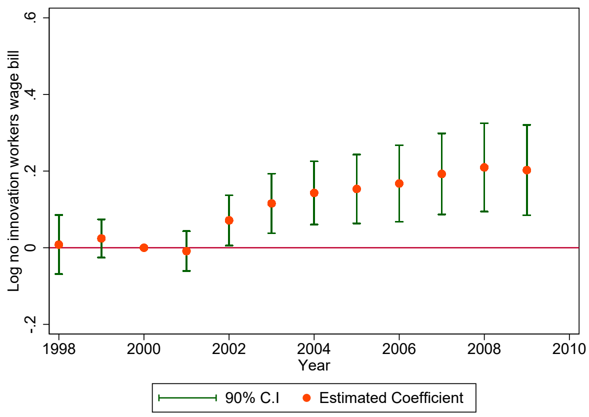
for $i = 1, 2, 3, 4$ where γ_i is a dummy for the different quartiles of R&D intensity. The dependent variable corresponds to the change in industry i share of aggregate employment in microregion k between 2000 and 2010. We split manufacturing industries in quartiles according to their level of R&D intensity so that 25% of the Brazilian manufacturing employment is in each group. Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC)

Figure 6: Effect of agricultural technical change on expenditure on non-innovative and innovative occupations

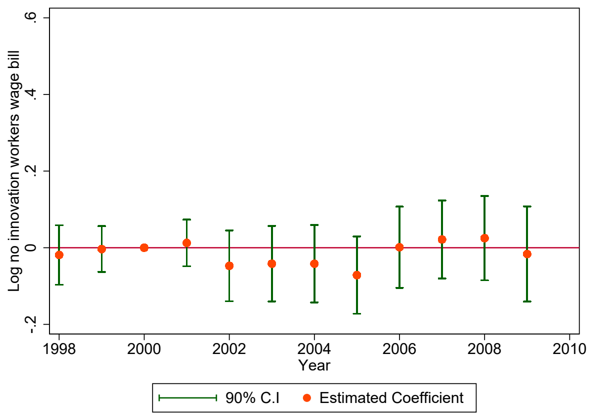
Yearly Social Security Data (1998-2009)

Low R&D Industries **High R&D Industries**

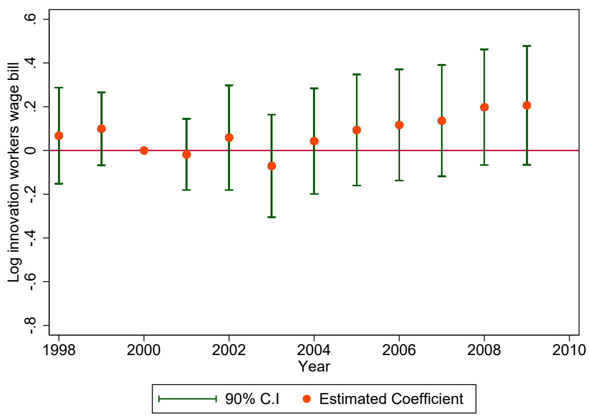
(a) **Non-innovative labor wage bill**



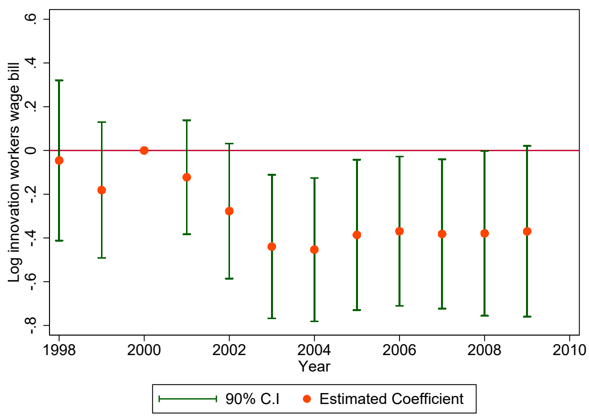
(b) **Non-innovative labor wage bill**



(c) **Innovative labor wage bill**



(d) **Innovative labor wage bill**



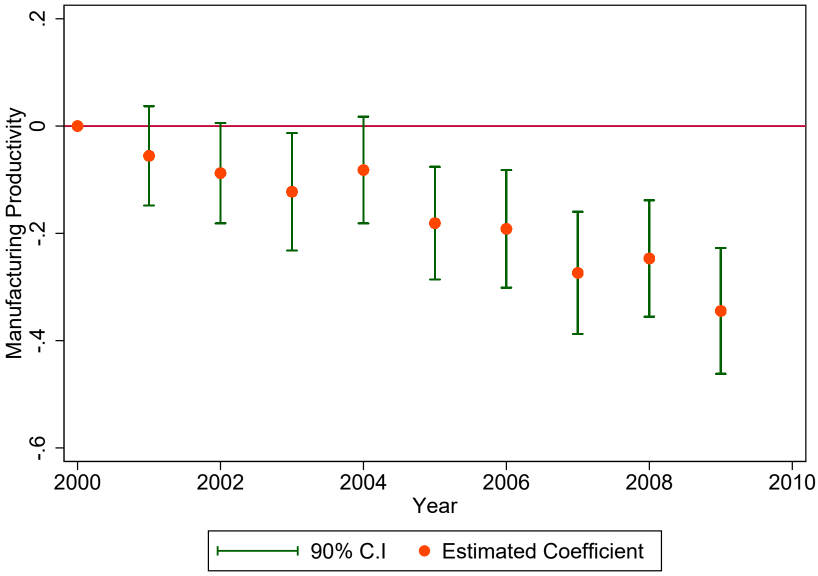
Notes: The plot shows the point estimates and the 90% confidence intervals for the estimates of the β_j coefficients of the following regression:

$$\ln y_{k,r,t} = \delta_t + \delta_k + \delta_{rt} + \sum_{j=1998}^{j=2009} \beta_j \Delta A_k^{soy} + tX'_{k,1991}\omega + \varepsilon_{k,t}$$

Standard errors are clustered at the microregion level. $\ln y_{k,t}$ corresponds to the log. wage bill on non-innovative and innovative labor in microregion k located in region r at the end of year t for Low R&D and High R&D manufacturing industries (Source: RAIS). Manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC).

Figure 7: Effect of agricultural technical change on manufacturing productivity

Yearly Manufacturing Survey Data (2000-2009)



Notes: The plot shows the point estimates and the 90% confidence intervals for the estimates of the β_j coefficients of the following regression:

$$\ln y_{k,r,t} = \delta_t + \delta_k + \delta_{rt} + \sum_{j=2000}^{j=2009} \beta_j \Delta A_k^{soy} + tX'_{k,1991}\omega + \varepsilon_{k,t}$$

The measure of manufacturing productivity, $\ln y_{k,r,t}$, corresponds to the aggregate log. value added per wage bill in microregion k located in region r at the end of year t for manufacturing industries. Standard errors are clustered at the microregion level. (Source: PIA).

Table 1: Effect of Agricultural Technical Change on GE Soy Adoption

Panel A	Δ GE-soy area share (1)	Δ GE-soy area share (2)	Δ non-GE soy area share (3)	Δ non-GE soy area share (4)
ΔA^{soy}	0.022***	0.020***	-0.007*	-0.008**
	[0.005]	[0.004]	[0.004]	[0.004]
Share rural population	0.034***	0.117***	-0.009	-0.057**
	[0.010]	[0.023]	[0.009]	[0.023]
Log Income per capita		-0.009		-0.002
		[0.006]		[0.007]
Literacy rate		0.162***		-0.043
		[0.034]		[0.035]
Log population density		0.005***		-0.006***
		[0.001]		[0.001]
Observations	557	557	557	557
R-squared	0.094	0.208	0.013	0.053

Panel B	Δ Soy area share (1)	Δ Soy area share (2)	Δ Maize area share (3)	Δ Maize area share (4)
ΔA^{soy}	0.022***	0.016***	-0.006	0.000
	[0.004]	[0.004]	[0.004]	[0.004]
ΔA^{maize}	-0.003**	-0.001	0.005***	0.003*
	[0.001]	[0.001]	[0.002]	[0.002]
Share rural population	0.029***	0.064***	0.020***	0.012
	[0.007]	[0.013]	[0.008]	[0.015]
Log Income per capita		-0.010*		-0.011
		[0.006]		[0.007]
Literacy rate		0.122***		-0.002
		[0.018]		[0.023]
Log population density		-0.001		0.003***
		[0.001]		[0.001]
Observations	557	557	556	556
R-squared	0.135	0.245	0.041	0.066

Notes: Changes in dependent variables are calculated over the years 1996 and 2006 (source: Agricultural Census). The unit of observation is the micro-region. Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 2: Summary Statistics of the Sample of Individuals by Sector

	2000	2010
Agriculture		
Age	38.0	39.0
Male (% of the Total)	89.3	81.2
White (% of the Total)	55.4	48.6
Education level (highest degree obtained)		
Less than Middle School (% of the Total)	86.1	72.7
Completed Middle School (% of the Total)	7.4	13.8
High School Graduates (% of the Total)	5.2	11.4
University Graduates (% of the Total)	1.3	2.1
Average log real hourly wage	0.81	1.06
For skilled labor	1.39	1.38
For unskilled labor	0.71	0.95
Low-R&D Manufacturing		
Age	36.8	37.3
Male (% of the Total)	61.6	58.7
White (% of the Total)	65.0	55.6
Education level (highest degree obtained)		
Less than Middle School (% of the Total)	52.2	36.8
Completed Middle School (% of the Total)	20.4	21.5
High School Graduates (% of the Total)	21.9	35.2
University Graduates (% of the Total)	5.5	6.6
Average log real hourly wage	1.23	1.51
For skilled labor	1.73	1.63
For unskilled labor	1.15	1.23
High-R&D Manufacturing		
Age	36.28	36.9
Male (% of the Total)	80.6	76.2
White (% of the Total)	63.0	55.2
Education level (highest degree obtained)		
Less than Middle School (% of the Total)	49.8	31.3
Completed Middle School (% of the Total)	20.0	19.8
High School Graduates (% of the Total)	23.4	39.8
University Graduates (% of the Total)	6.8	9.1
Average log real hourly wage	1.58	1.66
For skilled labor	1.92	1.81
For unskilled labor	1.24	1.35
Services		
Age	37.1	37.8
Male (% of the Total)	67.3	62.1
White (% of the Total)	58.9	50.8
Education level (highest degree obtained)		
Less than Middle School (% of the Total)	51.1	36.0
Completed Middle School (% of the Total)	17.9	19.3
High School Graduates (% of the Total)	23.4	34.3
University Graduates (% of the Total)	7.6	10.4
Average log real hourly wage	1.42	1.51
For skilled labor	1.77	1.67
For unskilled labor	1.01	1.24

Notes: The data comes from the Population Censuses for years 2000 and 2010. These summary statistics come from our final sample of individuals as detailed in Section 3.3. An individual is classified as skilled if it has at least completed the 8^{th} grade. This level should be attained when an individual is 14 or 15 years old and is equivalent to graduating from middle school. Manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC).

Table 3: Summary Statistics of the Sample of Microregions

		2000		Δ2000-2010		
	Source:	**Mean**	**SD**	**Mean**	**SD**	**Observations**
Potential Yields	*FAO-GAEZ*					
Soy		0.286	0.135	1.787	0.740	557
Maize		1.847	0.9984	3.082	1.639	557
Employment Shares	*Population Census*					
Agriculture		0.279	0.140	-0.050	0.055	557
Low-R&D Manufacturing		0.081	0.055	0.007	0.033	557
High-R&D Manufacturing		0.067	0.043	-0.001	0.025	557
Services		0.573	0.118	0.044	0.057	557
Log. Employment	*Population Census*					
Agriculture		8.268	0.890	0.122	0.249	557
Low-R&D Manufacturing		7.076	1.569	0.358	0.400	557
High-R&D Manufacturing		6.897	1.485	0.309	0.394	557
Services		9.194	1.887	0.404	0.175	557

	Source:	**Mean**	**SD**	**Observations**
Manufacturing Employment	*RAIS*			
Log. Employment				
Low-R&D Manufacturing		7.753	1.315	3,816
High-R&D Manufacturing		7.509	1.384	3,816
Log. Non-Innovative Labor				
Low-R&D Manufacturing		7.733	1.310	3,816
High-R&D Manufacturing		7.484	1.375	3,816
Log. Innovative Labor				
Low-R&D Manufacturing		3.530	1.857	3,816
High-R&D Manufacturing		3.308	2.214	3,816
Log. Non-Innovative Wage Bill				
Low-R&D Manufacturing		16.103	2.206	3,816
High-R&D Manufacturing		15.883	2.304	3,816
Log. Innovative Wage Bill				
Low-R&D Manufacturing		12.781	2.869	3,816
High-R&D Manufacturing		12.523	3.273	3,816
Manufacturing Productivity	*PIA*			
Log. Value Added per Worker				
Low-R&D Manufacturing		10.692	0.866	3,070
High-R&D Manufacturing		10.536	0.944	3,070
Log. Value Added per Wage Bill				
Low-R&D Manufacturing		1.537	0.593	3,070
High-R&D Manufacturing		1.360	0.613	3,070

Notes: The data sources are the Population Census (2000, 2010), RAIS and PIA. Manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC). A worker is classified as skilled if she has completed at least the 8^{th} grade (completed middle school).

Table 4: Effect of agricultural technical change on sectoral employment shares

Decadal Population Census Data (2000-2010)

Panel A

Outcome:	Change in employment shares by sector					
Sector:	Agriculture (1)	Agriculture (2)	Manufacturing (3)	Manufacturing (4)	Services (5)	Services (6)
ΔA_{soy}	-0.032***	-0.033***	0.025***	0.025***	0.007	0.008*
	[0.005]	[0.005]	[0.005]	[0.005]	[0.004]	[0.004]
Observations	557	557	557	557	557	557
R-squared	0.233	0.246	0.164	0.166	0.346	0.359
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes

Panel B

Outcome:	Change in employment shares of unskilled workers by sector					
Sector:	Agriculture (1)	Agriculture (2)	Manufacturing (3)	Manufacturing (4)	Services (5)	Services (6)
ΔA_{soy}	-0.033***	-0.033***	0.030***	0.030***	0.003	0.004
	[0.006]	[0.006]	[0.005]	[0.005]	[0.004]	[0.004]
Observations	557	557	557	557	557	557
R-squared	0.122	0.126	0.146	0.157	0.205	0.208
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes

Panel C

Outcome:	Change in employment shares of skilled workers by sector					
Sector:	Agriculture (1)	Agriculture (2)	Manufacturing (3)	Manufacturing (4)	Services (5)	Services (6)
ΔA_{soy}	-0.015***	-0.015***	0.014***	0.014***	0.001	0.001
	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.005]
Observations	557	557	557	557	557	557
R-squared	0.041	0.047	0.107	0.112	0.093	0.103
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). The unit of observation is the micro-region. All the regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 5: Effect of agricultural technical change on industrial specialization within manufacturing

Decadal Population Census Data (2000-2010)

Panel A

Outcome:	Change in employment shares by manufacturing industry		
Industry:	All	Low R&D Intensive	High R&D Intensive
	(1)	(2)	(3)
ΔA_{soy}	0.025***	0.021***	0.004**
	[0.005]	[0.004]	[0.002]
Observations	557	557	557
R-squared	0.166	0.135	0.172
Baseline Controls	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
All Controls	Yes	Yes	Yes

Panel B

Outcome:	Change in employment shares of unskilled workers by manufacturing industry		
Industry:	All	Low R&D Intensive	High R&D Intensive
	(1)	(2)	(3)
ΔA_{soy}	0.030***	0.026***	0.004**
	[0.005]	[0.004]	[0.002]
Observations	557	557	557
R-squared	0.157	0.159	0.101
Baseline Controls	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
All Controls	Yes	Yes	Yes

Panel C

Outcome:	Change in employment shares of skilled workers by manufacturing industry		
Industry:	All	Low R&D Intensive	High R&D Intensive
	(1)	(2)	(3)
ΔA_{soy}	0.014***	0.012**	0.002
	[0.005]	[0.005]	[0.002]
Observations	557	557	557
R-squared	0.112	0.080	0.080
Baseline Controls	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
All Controls	Yes	Yes	Yes

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). The unit of observation is the micro-region. All the regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. In these regressions, manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC). Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 6: Effect of agricultural technical change on industry size and innovation expenditures

Yearly Social Security Data (1998-2009)

	Industry Size		Innovation Expenditures	
Outcomes:	Wage Bill of Non-Innovation Workers		Wage Bill of Innovation Workers	
Industry:	Low R&D (1)	High R&D (2)	Low R&D (3)	High R&D (4)
A_{soy}	0.148***	-0.008	0.060	-0.274*
	[0.047]	[0.046]	[0.109]	[0.152]
Observations	3,816	3,816	3,816	3,816
R-squared	0.984	0.988	0.940	0.929
Baseline Controls	Yes	Yes	Yes	Yes
Region x Year FEs	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes

Notes: The dependent variables correspond to the wage bill on non-innovation workers (in logs) for each manufacturing industry in each microregion as a proxy for industry size and the wage bill on innovation workers (in logs) for each type of industry in every microregion as a proxy for expenditure on innovation. We use aggregate information from RAIS at the microregion-industry level for the time period 1998-2009. We include only those microregions that have positive employment for all the years in the sample. A^{soy} is defined as potential soy yield under high inputs for the years between 2003 and 2009, and the potential soy yield under low inputs for the years between 1998 and 2002. Baseline controls include the share of rural population in 1991, a measure of technical change in maize and region-year fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in 1991, all interacted with a linear trend. The unit of observation is a microregion. In these regressions, manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC). Standard errors clustered at the microregion level reported in parentheses. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 7: Effect of agricultural technical change on manufacturing productivity

Yearly Manufacturing Survey Data (2000-2009)

Panel A

Outcomes:	Log Value Added per Wage Bill			High R&D Innovation Share measured by wage bill
Industry:	All	Low R&D	High R&D	All
	(1)	(2)	(3)	(4)
A_{soy}	-0.133***	-0.135**	-0.109*	-0.043**
	[0.040]	[0.054]	[0.057]	[0.020]
Observations	3,070	3,070	3,070	3,816
R-squared	0.735	0.627	0.635	0.755
Baseline Controls	Yes	Yes	Yes	Yes
Region x Year FEs	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes

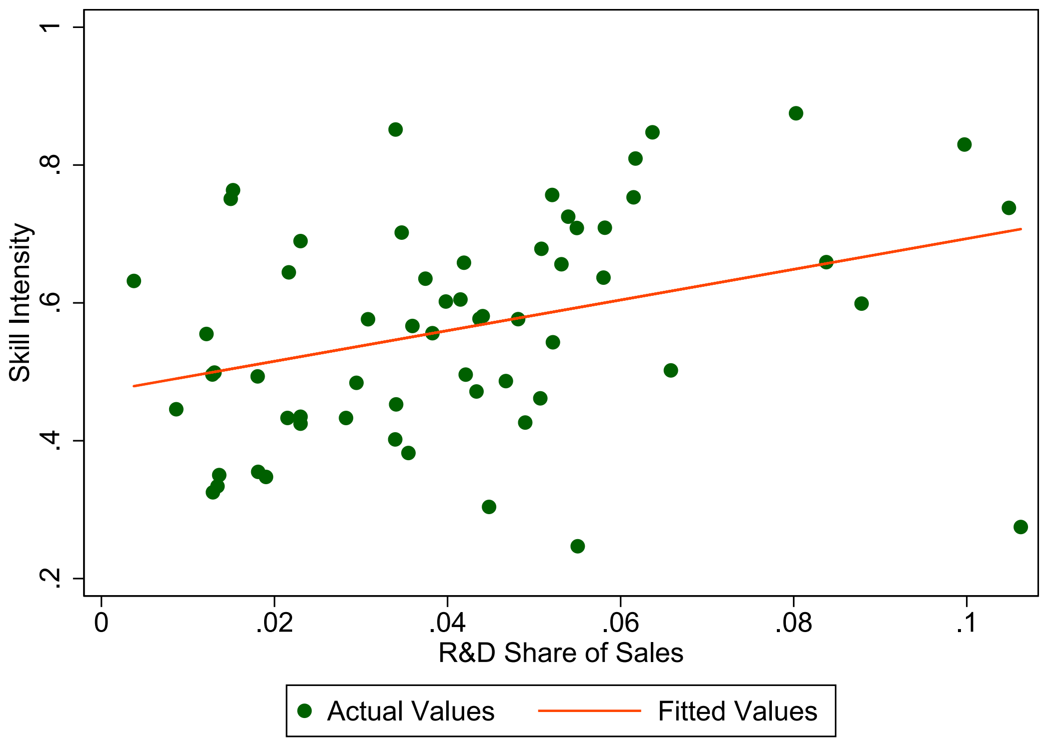
Panel B

Outcomes:	Log Value Added per Worker			High R&D Innovation Share measured by employment
Industry:	All	Low R&D	High R&D	All
	(1)	(2)	(3)	(4)
A_{soy}	-0.141***	-0.151**	-0.119*	-0.051***
	[0.043]	[0.059]	[0.071]	[0.018]
Observations	3,070	3,070	3,070	3,816
R-squared	0.876	0.796	0.799	0.747
Baseline Controls	Yes	Yes	Yes	Yes
Region x Year FEs	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes

Notes: The dependent variables correspond to the total value added divided by total wage bill (in logs) and total value added divided by employment (in logs) for each type of manufacturing industry in each microregion as a proxy for productivity and the share of the High R&D industry innovation share as measured by the share of the innovation workers' wage bill working in the High R&D industry and share of the innovation workers working in the High R&D industry. We use aggregate information from PIA at the microregion level for the time period 2000-2009 for Columns (1)-(3) and aggregate information from RAIS at the microregion-industry level for the time period 1998-2009 for Column (4). We include only those microregions that have positive employment for all the years in the sample. A_{soy} is defined as potential soy yield under high inputs for the years between 2003 and 2009, and the potential soy yield under low inputs for the years between 2000 and 2002. Baseline controls include the share of rural population in 1991, a measure of technical change in maize and region year fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in 1991, all interacted with a linear trend. The unit of observation is a microregion. In these regressions, manufacturing industries are classified as Low-R&D or High-R&D intensive depending on whether their R&D intensity is below or above the median in 2000 (weighting industries by number of employees so that each group captures around 50 percent of total manufacturing employment). We define R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC). Standard errors clustered at the microregion level reported in parentheses. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

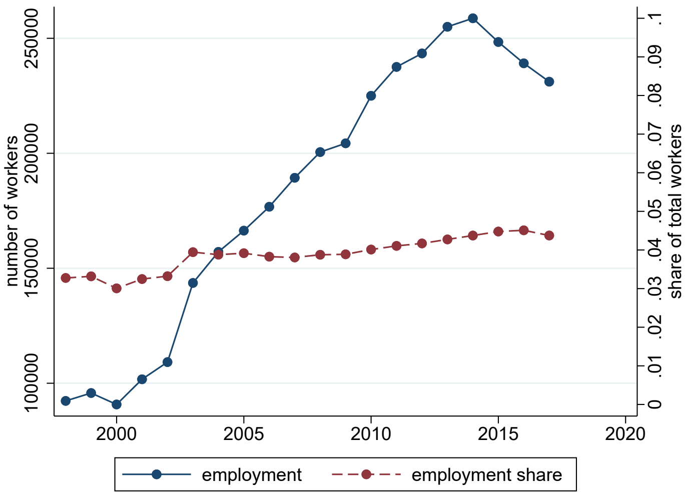
A Appendix: Figures and Tables

Figure A.1: Correlation between Skill Intensity and R&D Intensity at Industry Level



Notes: We define skill intensity as the share of skilled individuals in a particular industry in Brazil at baseline and we source it from the 2000 Population Census. Our measure of R&D activity is R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica](PINTEC)*. The correlation between these variables is approximately 0.34.

Figure A.2: Manufacturing Employment in Innovation Intensive Occupations



Notes: Authors' calculations using RAIS data. Innovation intensive occupations are defined using the methodology described in section 3.3.

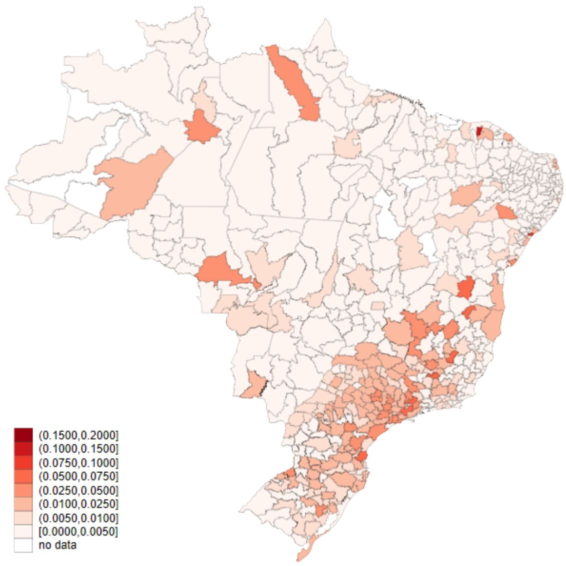
Figure A.3: Correlations between share of workers in innovation intensive occupations and other industry-level measures of innovation

(a) R&D expenditure per worker

(b) Share of High-Skill workers

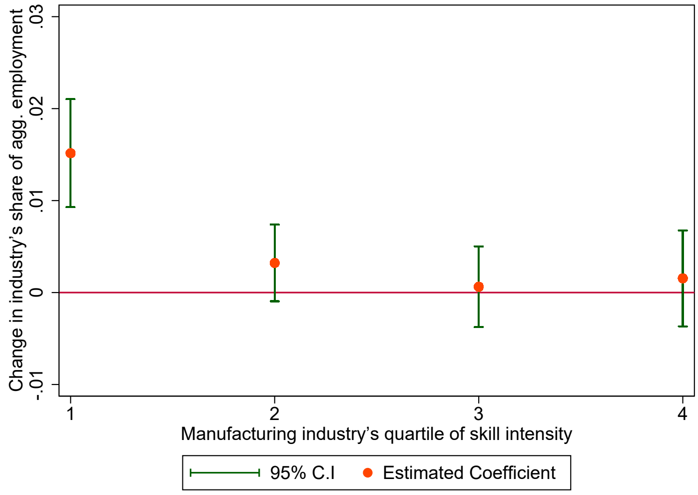
Notes: Innovation intensive occupations are defined using the methodology described in section 3.3. The share of workers in innovation intensive occupations in each sector is constructed using RAIS data for the year 2000. Skill intensity is the share of skilled individuals in each sector constructed using RAIS data for the year 2000. R&D expenditure per worker is defined as R&D expenditure from the 2000 *Pesquisa de Inovação Tecnológica](PINTEC)* divided by number of workers in each industry in 2000 (from RAIS).

Figure A.4: Geographical distribution of share of manufacturing workers in innovation intensive occupations in 2000



Notes: Authors' calculations using RAIS data and Population Census data for year 2000. Innovation intensive occupations are defined using the methodology described in section 3.3. The Figure reports the share of innovation intensive workers over total workers in the manufacturing sector in the year 2000 by microregion.

Figure A.5: Employment Share Growth by Quartile of Skill Intensity



Notes: The plot shows the β_i coefficients of the following regression:

$$\Delta \frac{L^k_{m,i}}{L^k} = \alpha + \beta_i \Delta A_{soy} \times \gamma_i + \gamma_i + \varphi X_{k,1991} + \varepsilon^i_k$$

for $i = 1, 2, 3, 4$ where γ_i is a dummy for the different quartiles of skill intensity. We split manufacturing industries in quartiles according to their level of skill and R&D intensity so that 25% of the Brazilian manufacturing employment is in each group. Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). We define skill intensity as the share of skilled individuals in a particular industry in Brazil at baseline and we source it from the 2000 Population Census.

Table A.1: Classification of Manufacturing Industries by R&D Intensity

IBGE Code	Description	R&D Share of Sales	Skill Intensity
26091	Ceramic products	0.106	0.275
34001	Manufacturing and assembly of motor vehicles	0.105	0.738
23030	Production of nuclear fuels	0.100	0.830
31002	Electrical material for vehicles	0.088	0.599
27001	Steel products	0.084	0.659
35030	Construction, assembly and repair of airplanes	0.080	0.875
28002	Foundries, stamping shops, powder metallurgy and metal treatment services	0.066	0.502
33003	Machines, equipment for electronic systems for industrial automation, and control	0.064	0.848
24020	Pharmaceutical products	0.062	0.809
33001	Medical equipment	0.061	0.753
29002	Appliances	0.058	0.709
34002	Cabins, car bodies, trailers and parts for motor vehicles	0.058	0.637
20000	Wooden products	0.055	0.247
33004	Equipment, instruments and optical, photographic and cinematographic material	0.055	0.709
33002	Measuring, testing and control equipment - except for controlling industrial processes	0.054	0.725
24010	Paints, dyes, varnish, enamels and lacquers	0.053	0.656
25020	Plastic products	0.052	0.543
32000	Electronic material and communications equipment	0.052	0.757
31001	Machines, equipment and miscellaneous electric material - except for vehicles	0.051	0.678
27003	Foundries	0.051	0.462
15043	Other food products	0.049	0.426
36090	Miscellaneous products	0.048	0.576
23010	Coke plants	0.047	0.487
37000	Recycling	0.045	0.304
35090	Miscellaneous transportation equipment	0.044	0.581
21002	Corrugated cardboard, packaging, and paper and cardboard objects	0.044	0.577
17001	Processing of fibers, weaving and cloth making	0.043	0.471
28001	Metal products - except machines and equipment	0.042	0.496
24030	Soap, detergents, cleaning products and toiletries	0.042	0.658
29001	Machines and equipment - except appliances	0.041	0.605
21001	Pulp, paper and smooth cardboard, poster paper and card paper	0.040	0.602
34003	Reconditioning or restoration of engines of motor vehicles	0.038	0.556
24090	Miscellaneous chemical products	0.037	0.635
25010	Rubber products	0.036	0.567
26092	Miscellaneous products of non-metallic minerals	0.035	0.382
22000	Editing, printing and reproduction of recordings	0.035	0.702
19012	Leather objects	0.034	0.453
30000	Office machines and data-processing equipment	0.034	0.852
36010	Pieces of furniture	0.034	0.402
26010	Glass and glass products	0.031	0.576
15021	Preserves of fruit, vegetables and other vegetable products	0.029	0.484
17002	Manufacturing of textile objects based on cloth - except for garments	0.028	0.433
18001	Making of clothing articles and accessories - except on order	0.023	0.425
18002	Making clothing articles and accessories - on order	0.023	0.435
18999	Making of clothing articles and accessories - on order or not	0.023	0.690
27002	Non-ferrous metals	0.022	0.644
15030	Dairy products	0.022	0.433
19020	Footwear	0.019	0.348
15010	Slaughtering and preparation of meat and fish	0.018	0.355
35010	Construction and repair of boats	0.018	0.493
23020	Products in oil refining	0.015	0.763
33005	Chronometers, clocks and watches	0.015	0.751
23400	Alcohol production	0.014	0.350
15041	Manufacturing and refining of sugar	0.013	0.334
15042	Roasting and grinding of coffee	0.013	0.499
19011	Tanning and other preparations of leather	0.013	0.325
16000	Tobacco products	0.013	0.496
15050	Beverages	0.012	0.555
15022	Vegetable fat and oil	0.009	0.446
35020	Construction and assembly of locomotives, cars and other rolling stock	0.004	0.632
	Median	0.041	0.432

Notes: The industry codes correspond to the CNAE-Domiciliar, the industry classification used in the 2000 Population Census. Industries are sorted by their R&D intensity at baseline. We measure R&D intensity as R&D expenditure as a share of total sales at baseline and we source it from from the 2000 *Pesquisa de Inovação Tecnológica* (PINTEC). We define skill intensity as the share of skilled individuals in a particular industry in Brazil at baseline and we source it from the 2000 Population Census. The correlation between these variables is approximately 0.34. Industries below the median are classified as low and the ones above the median as high.

Table A.2: Keywords Identifying Innovative Occupations

Portuguese	English
Nouns or combination of nouns in task description of occupations	
pesquisa e desenvolvimento	research and development
inovação	innovation
p&d	R&D
desenvolvimento de produtos	product development
desenvolvimento de processos	process development
pesquisador	researcher
novas tecnologias	new technologies
protótipos	prototypes
pesquisas tecnologicas	technological research
automaçao de processos	process automation
Actions (verb + noun) in task description of occupations	
desenvolvem produtos	develop products
desenvolvem pesquisas	develop research
desenvolvem equipamentos	develop equipment
desenvolvem processos	develop processes
desenvolvem dispositivos	develop devices
otimizam métodos	optimize methods
otimizam os meios	optimize means
aperfeiçoam sistemas	improve systems
aperfeiçoam processos	improve processes
aperfeiçoam produtos	improve products
aperfeiçoam dispositivos	improve devices
implementam dispositivos de automaçao	implement automation devices
desenvolvem, testam e supervisionam sistemas, processos e métodos produtivos	develop, test and supervise systems, processes and production methods
Nouns or combinations of nouns (source: Technical Appendix of the 2008 PINTEC survey)	
produto novo / novo produto	new product
produtos novos / novos produtos	new products
produto aprimorado	improved product
produtos aprimorados	improved produts
inovação de produto	product innovation
aperfeiçoamento de produto	product improvement
processo novo / novo processo	new process
processos novos / novos processos	new processes
processo aprimorado	improved process
processos aprimorados	improved processes
inovação de processo	process innovation
aperfeiçoamento de processo	process improvement

Notes: The Table reports the keywords used to identify innovation intensive occupations and their English translation. Task descriptions for each occupations are obtained from the official publication of the "Brazilian Classification of Occupations", Ministry of Labor, 3rd Edition (2010).

Table A.3: Internal migration

Outcomes	$\Delta \log L$	Net Migration	In-Migration	Out-Migration	Net Migration	In-Migration	Out-Migration	Net Migration	In-Migration	Out-Migration
Skill Group:		All				Skilled			Unskilled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ΔA_{soy}	-0.014	0.004	0.002	-0.003	-0.001	-0.004	-0.003	0.012	0.011**	-0.002
	[.0134]	[0.009]	[0.005]	[0.006]	[0.010]	[0.005]	[0.007]	[0.008]	[0.005]	[0.006]
Observations	557	557	557	557	557	557	557	557	557	557
R-squared	0.171	0.553	0.401	0.592	0.507	0.380	0.593	0.582	0.407	0.566
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variables are calculated for 2010 (source: Population Censuses). The unit of observation is the micro-region. These regressions compute the 5 year internal migration rate between 2005 and 2010, using the microregion of residence 5 years prior to the Census 2010. All the regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. . Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table A.4: Effect of technical change in soy on wages by sector

Outcome:	Change in composition-adjusted wages by sector							
Sector	Overall (1)	Overall (2)	Agriculture (3)	Agriculture (4)	Manufacturing (5)	Manufacturing (6)	Services (7)	Services (8)
ΔA_{soy}	0.012	0.023***	0.042***	0.051***	0.009	0.018	0.004	0.016*
	[0.008]	[0.007]	[0.010]	[0.010]	[0.011]	[0.011]	[0.009]	[0.008]
Observations	557	557	557	557	557	557	557	557
R-squared	0.241	0.355	0.319	0.374	0.082	0.139	0.191	0.319
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Changes in wages are calculated over the years 2000 to 2010. The unit of observation is the micro-region. All the regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. We recover the estimates of the dependent variable from a first stage Mincerian regression in which we estimate a regression of the log of hourly wage on microregion fixed effects, and a vector of individual characteristics that includes dummies for sector, skill group, age group, race, and all the interactions between these variables. Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A.5: Effect of technical change in soy on wages by skill group

Panel A

Outcome:	Change in composition-adjusted wages of unskilled workers by sector							
Sector	Overall (1)	Overall (2)	Agriculture (3)	Agriculture (4)	Manufacturing (5)	Manufacturing (6)	Services (7)	Services (8)
ΔA_{soy}	0.003	0.012	0.039***	0.045***	0.003	0.012	-0.002	0.010
	[0.008]	[0.008]	[0.011]	[0.012]	[0.013]	[0.012]	[0.009]	[0.009]
Observations	557	557	557	557	556	556	557	557
R-squared	0.339	0.387	0.323	0.170	0.060	0.104	0.185	0.293
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel B

Outcome:	Change in composition-adjusted wages of skilled workers by sector							
Sector	Overall (1)	Overall (2)	Agriculture (3)	Agriculture (4)	Manufacturing (5)	Manufacturing (6)	Services (7)	Services (8)
ΔA_{soy}	0.024**	0.033***	0.055***	0.065***	0.029	0.042**	0.022**	0.034***
	[0.010]	[0.010]	[0.019]	[0.018]	[0.018]	[0.017]	[0.011]	[0.011]
Observations	557	557	557	557	555	555	557	557
R-squared	0.157	0.216	0.179	0.199	0.063	0.107	0.132	0.217
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes	No	Yes

Panel C

Outcome:	Change in composition-adjusted skill premia by sector							
Sector	Overall (1)	Overall (2)	Agriculture (3)	Agriculture (4)	Manufacturing (5)	Manufacturing (6)	Services (7)	Services (8)
ΔA_{soy}	0.021**	0.021**	0.016	0.017	0.026	0.028	0.024**	0.024**
	[0.009]	[0.009]	[0.019]	[0.019]	[0.020]	[0.019]	[0.010]	[0.010]
Observations	557	557	557	557	554	554	557	557
R-squared	0.162	0.165	0.145	0.150	0.027	0.042	0.031	0.032
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Changes in wages and skill premia are calculated over the years 2000 to 2010. All regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. In columns (5) and (6) of Panel A we lose one observation because there are no unskilled manufacturing workers in our sample in the microregion Amapá (IBGE ID 16002) in 2010. In columns (5) and (6) of Panel B we lose two observations because there are no skilled male manufacturing workers in our sample in the microregions of Japurà (IBGE ID 13002) and Chapadas Das Mangabeiras (IBGE ID 21021) in 2000. The missing observations in columns (5) and (6) of Panel C follow from the above explanation. We recover the estimates of the dependent variable from a first stage Mincerian regression in which we estimate a regression of the log of hourly wage on microregion fixed effects, and a vector of individual characteristics that includes dummies for sector, skill group, age group, race, and all the interactions between these variables. Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A.6: Effect of technical change in soy on the number of workers at the minimum wage

Outcome:	Change in the share of workers at the minimum wage by sector					
Industry	Manufacturing	Manufacturing	Low R&D	Low R&D	High R&D	High R&D
ΔA_{soy}	0.210***	0.177***	0.245***	0.207***	0.231***	0.185***
	[0.045]	[0.044]	[0.052]	[0.048]	[0.048]	[0.045]
Observations	556	556	555	555	555	555
R-squared	0.124	0.184	0.146	0.221	0.213	0.302
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	No	Yes	No	Yes	No	Yes

Notes: Changes in dependent variables are calculated over the years 2000 and 2010 (source: Population Censuses). The unit of observation is the micro-region. Workers at the minimum wage are workers paid below the mandatory minimum wage in 2000 and 2010. All the regressions include the baseline specification controls which are the share of rural population in 1991, a measure of technical change in maize and region fixed effects. The regressions with all controls also include income per capita (in logs), population density (in logs), literacy rate, all observed in the 1991 Population Census. Robust standard errors reported in brackets. Significance levels: $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$..

B Appendix: data

B.1 Wages

To compute composition-adjusted wages we estimate the following Mincerian regressions:

$$ln(w_{ikt}) = \gamma_{kt} + H_{ikt}\beta_{Ht} + \varepsilon_{ikt} \quad \textit{for t=2000, 2010} \tag{16}$$

where $ln(w_{ijkt})$ is the log hourly wage of individual i, working in sector j in micro-region k at time t, and γ_{kt} is a micro-region fixed effect, while H_{ijkt} is a vector of individual characteristics, which includes dummies for sector, skill group, age group, race, and all the interactions between these variables. We estimate the previous Mincerian regression for each micro-region and for each broad sector separately. Also, we estimate these regressions constraining the sample to either unskilled or skilled labor only, recovering the unit price of labor in each micro-region for each type of labor in both cross sections. Since the existing literature documented how Brazil has experienced a considerable reduction in its gender pay gap (Ferreira, Firpo, and Messina 2017), we estimate equation (16) only for male workers. Observations are weighted by their corresponding population census weight. Next, we use the micro-region fixed effects estimated above as the unit price of labor for a given skill group in a given micro-region, and we compute the change in unit prices of labor in micro-region k between 2000 and 2010 as $\Delta\gamma_k = \gamma_{k,2010} - \gamma_{k,2000}$, which gives us the change in the composition-adjusted wages at the micro-region level.

C Appendix: Theory

In this appendix we provide the proofs of Propositions 1 to 4 and Lemma 1.

Proposition 1. *An increase in A_s in agriculture, leads to an increase in the relative demand for high skilled workers in agriculture if and only if the elasticity of substitution between high- and low-skilled workers is greater than one ($\varepsilon > 1$).*

Proof. Take the agriculture sector. Solving for the inner nest we get that the conditional factor demands $S_a(w_s, w_u, L_a)$, $U_a(w_s, w_u, L_a)$ and the cost function $C(w_s, w_u, L_a)$ for agriculture labor L_a are given by:

$$S_a(w_s, w_u, L_a) = \frac{\left(\frac{w_s}{A_s}\right)^{-\varepsilon} L_a}{A_s \left[w_s^{1-\varepsilon} A_s^{\varepsilon-1} + w_u^{1-\varepsilon} A_u^{\varepsilon-1}\right]^{\frac{\varepsilon}{\varepsilon-1}}} \tag{17}$$

$$U_a(w_s, w_u, L_a) = \frac{\left(\frac{w_u}{A_u}\right)^{-\varepsilon} L_a}{A_u \left[w_s^{1-\varepsilon} A_s^{\varepsilon-1} + w_u^{1-\varepsilon} A_u^{\varepsilon-1}\right]^{\frac{\varepsilon}{\varepsilon-1}}} \tag{18}$$

$$C(w_s, w_u, L_a) = L_a \left[\left(\frac{w_s}{A_s}\right)^{1-\varepsilon} + \left(\frac{w_u}{A_u}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} \tag{19}$$

Thus, the relative demand for skilled workers in agriculture is given by:

$$\frac{S_a}{U_a} = \left(\frac{w_u}{w_s}\right)^{\varepsilon} \left(\frac{A_s}{A_u}\right)^{\varepsilon-1} \tag{20}$$

□

Proposition 2. *Whether an increase in A_s in agriculture leads to an absolute decrease in the demand for low skilled workers in agriculture depends on whether labor and land are strong complements ($\sigma < \varepsilon\Gamma$).*

Proof. From the production function we can can compute the marginal productivity for each raw labor type:

$$MPU_a = K_t^h \gamma \Theta^{\frac{1}{\sigma-1}} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}} A_u^{\frac{\varepsilon-1}{\varepsilon}} U_a^{\frac{-1}{\varepsilon}} \tag{21}$$

$$MPS_a = K_t^h \gamma \Theta^{\frac{1}{\sigma-1}} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}} A_s^{\frac{\varepsilon-1}{\varepsilon}} S_a^{\frac{-1}{\varepsilon}} \tag{22}$$

where $\Theta = (A_L L_a)^{\frac{\sigma-1}{\sigma}} + (A_T T_a)^{\frac{\sigma-1}{\sigma}}$. Clearly, we can see that

$$\frac{\partial \Theta}{\partial A_s} = \frac{\sigma-1}{\sigma} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{\sigma-\varepsilon}{\sigma\varepsilon}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}}$$

Moreover,

$$\frac{\partial L_a^m}{\partial A_s} = m L_a^{m-1+\frac{1}{\varepsilon}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}}$$

Therefore,

$$\frac{\partial MPU_a}{\partial A_s} = A_n K_t^h A_L^{\frac{\sigma-1}{\sigma}} A_u^{\frac{\varepsilon-1}{\varepsilon}} U_a^{\frac{-1}{\varepsilon}} \left(\frac{1}{\sigma-1} \Theta^{\frac{2-\sigma}{\sigma-1}} \frac{\partial \Theta}{\partial A_s} L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}} + \Theta^{\frac{1}{\sigma-1}} \frac{\partial L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}}}{\partial A_s} \right)$$

$$\frac{\partial MPU_a}{\partial A_s} = \underbrace{A_n K_t^h A_L^{\frac{\sigma-1}{\sigma}} A_u^{\frac{\varepsilon-1}{\varepsilon}} U_a^{\frac{-1}{\varepsilon}} \Theta^{\frac{1}{\sigma-1}} L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}}}_{\kappa} \left(\frac{1}{\sigma-1}\Theta^{-1}\frac{\partial\Theta}{\partial A_s} - \frac{(\varepsilon-\sigma)}{\varepsilon\sigma} L_a^{-1}\frac{\partial L_a}{\partial A_s} \right)$$

Notice that $\kappa > 0$. Thus,

$$\frac{\partial MPU_a}{\partial A_s} = \kappa \left(\frac{1}{\sigma}\Theta^{-1} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{\sigma-\varepsilon}{\sigma\varepsilon}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}} - \frac{(\varepsilon-\sigma)}{\varepsilon\sigma} L_a^{\frac{1-\varepsilon}{\varepsilon}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}} \right)$$

$$\frac{\partial MPU_a}{\partial A_s} = \frac{\kappa}{\sigma} L_a^{\frac{1}{\varepsilon-1}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}} \left(\Theta^{-1} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{\sigma-\varepsilon}{\sigma\varepsilon}} - \frac{(\varepsilon-\sigma)}{\varepsilon} L_a^{\frac{1-\varepsilon}{\varepsilon}} \right)$$

Since $\frac{\kappa}{\sigma} L_a^{\frac{1}{\varepsilon-1}} S_a^{\frac{\varepsilon-1}{\varepsilon}} A_s^{\frac{-1}{\varepsilon}} > 0$

$$\frac{\partial MPU_a}{\partial A_s} < 0 \iff \Theta^{-1} A_L^{\frac{\sigma-1}{\sigma}} L_a^{\frac{\sigma-\varepsilon}{\sigma\varepsilon}} - \frac{(\varepsilon-\sigma)}{\varepsilon} L_a^{\frac{1-\varepsilon}{\varepsilon}} < 0$$

$$\frac{\partial MPU_a}{\partial A_s} < 0 \iff \sigma < \varepsilon \left(\frac{(A_L L_a)^{\frac{\sigma-1}{\sigma}} + (A_T T_a)^{\frac{\sigma-1}{\sigma}} - (A_L L_a)^{\frac{\sigma-1}{\sigma}}}{\Theta} \right)$$

$$\frac{\partial MPU_a}{\partial A_s} < 0 \iff \sigma < \varepsilon \left(\frac{(A_T T_a)^{\frac{\sigma-1}{\sigma}}}{\Theta} \right) \tag{23}$$

□

Lemma 1. *If all three sectors are active, the effect of an increase in skilled-biased-factor-augmenting technology in agriculture (A_s) on wages is mediated by the effect of A_s on local knowledge (K_t^h). In particular:*

$$\frac{\partial \ln w_s}{\partial A_s} = \frac{\partial \ln w_u}{\partial A_s} = \frac{\partial \ln K_t^h}{\partial A_s}$$

and the effect of A_s on land prices is given by:

$$\frac{\partial \ln w_T}{\partial A_s} = \frac{\partial \ln K_t^h}{\partial A_s} + \frac{\theta_{S_a}}{A_s \theta_{T_a}}$$

where θ_{S_a} is the cost share of high-skilled workers and θ_{T_a} is the cost share of land in agriculture.

Proof. The unit cost functions are defined as:

$$c_a(w_s, w_u, w_T, A_s, K_t^h) = \min\{w_s S_a + w_u U_a + w_T T_a \mid Y_a \geq 1\}$$

$$c_m^h(w_s, w_u, w_T, p, K_t^h) = \min\{w_s S_m^h + w_u U_m^h + p^h K_t^h x^h \mid Y_m^h \geq 1\}$$

$$c_m^\ell(w_s, w_u, w_T, K_t) = \min\{w_s S_m^\ell + w_u U_m^\ell + p^\ell K_t^\ell x^\ell \mid Y_m^\ell \geq 1\}$$

Where A_s denotes skilled-biased factor-augmenting technologies in agriculture, K_t^h is the local knowledge which is an endogenous hicks neutral technology, and p^j is the price of inputs and x^j is the quantity of inputs. Note that we already use the symmetry of the input market to simplify notation.

From the unit cost functions we can define the unit factor demands:

$$a_{U_i}(w_s, w_u, w_T, A_i, K_t^h) = \frac{\partial c_i(w_s, w_u, w_T, A_i, K_t^h)}{\partial w_u}$$

$$a_{S_i}(w_s, w_u, w_T, K_t^h) = \frac{\partial c_i(w_s, w_u, w_T, K_t^h)}{\partial w_s}$$

$$a_{T_i}(w_s, w_u, w_T, K_t^h) = \frac{\partial c_i(w_s, w_u, w_T, K_t^h)}{\partial r}$$

In this economy, when all sectors are active, zero profit conditions are given by:

$$p_a = c_a(w_s, w_u, w_T, A_s, K_t^h) = c_a(w_s, w_u, w_T, A_s)/K_t^h$$

$$1 = c_m^h(w_s, w_u, p, K_t^h) = c_m^h(w_s, w_u, p^h)/K_t^h$$

$$p_m^\ell = c_m^\ell(w_s, w_u, K_t^h) = c_m^\ell(w_s, w_u, p^\ell)/K_t^h$$

These equations can be re-written as:

$$p_a = c_a(\frac{w_s}{A_s}, w_u, w_T)/K_t^h$$

$$1 = c_m^h(w_s, w_u, p^h)/K_t^h$$

$$p_m^\ell = c_m^\ell(w_s, w_u, p^\ell)/K_t^h$$

Where we made clear that the unit cost function in agriculture depends on the skilled biased factor-augmenting technology A_s that we study, and that the productivity in all sectors also depends on K_t. Taking log derivatives of these equations with respect to A_s we obtain that:

$$\frac{\partial \ln p_a}{\partial A_s} = \theta_{T_a}\frac{\partial \ln w_T}{\partial A_s} + \theta_{S_a}\frac{\partial \ln w_s}{\partial A_s} - \theta_{S_a}\frac{\partial \ln A_s}{\partial A_s} + \theta_{U_a}\frac{\partial \ln w_u}{\partial A_s} - \frac{\partial \ln K_t^h}{\partial A_s}$$

$$\frac{\partial \ln 1}{\partial A_s} = \theta_{S^h_m}\frac{\partial \ln w_s}{\partial A_s} + \theta_{U^h_m}\frac{\partial \ln w_u}{\partial A_s} + \theta_{x^h_m}\frac{\partial \ln p^h}{\partial A_s} - \frac{\partial \ln K^h_t}{\partial A_s}$$

$$\frac{\partial \ln p^\ell_m}{\partial A_s} = \theta_{S^\ell_m}\frac{\partial \ln w_s}{\partial A_s} + \theta_{U^\ell_m}\frac{\partial \ln w_u}{\partial A_s} + \theta_{x^\ell_m}\frac{\partial \ln p^\ell}{\partial A_s} - \frac{\partial \ln K^h_t}{\partial A_s}$$

But, we will later see that the price of inputs is proportional to the cost of producing them. And the cost of producing one input is the same as the final good.[37] Defining:

$$\tilde{\theta}_{S^j_m} = (\theta_{S^j_m} + \theta_{x^j_m}\frac{\theta_{S^j_m}}{\theta_{S^j_m} + \theta_{U^j_m}})$$

We then have:

$$\frac{\partial \ln K^h_t}{\partial A_s} = \tilde{\theta}_{S^h_m}\frac{\partial \ln w_s}{\partial A_s} + \tilde{\theta}_{U^h_m}\frac{\partial \ln w_u}{\partial A_s}$$

$$\frac{\partial \ln K^h_t}{\partial A_s} = \tilde{\theta}_{S^\ell_m}\frac{\partial \ln w_s}{\partial A_s} + \tilde{\theta}_{U^\ell_m}\frac{\partial \ln w_u}{\partial A_s}$$

Hence:

$$\frac{\partial \ln K^h_t}{\partial A_s} = \tilde{\theta}_{S^h_m}\frac{\partial \ln w_s}{\partial A_s} + (1-\tilde{\theta}_{S^h_m})\frac{\partial \ln w_u}{\partial A_s}$$

$$\frac{\partial \ln K^h_t}{\partial A_s} = \tilde{\theta}_{S^\ell_m}\frac{\partial \ln w_s}{\partial A_s} + (1-\tilde{\theta}_{S^\ell_m})\frac{\partial \ln w_u}{\partial A_s}$$

In matrix form:

$$\begin{bmatrix} \frac{\partial \ln K^h_t}{\partial A_s} \\ \frac{\partial \ln K^h_t}{\partial A_s} \end{bmatrix} = \begin{bmatrix} \tilde{\theta}_{S^h_m} & (1-\tilde{\theta}_{S^h_m}) \\ \tilde{\theta}_{S^\ell_m} & (1-\tilde{\theta}_{S^\ell_m}) \end{bmatrix} \begin{bmatrix} \frac{\partial \ln w_s}{\partial A_s} \\ \frac{\partial \ln w_u}{\partial A_s} \end{bmatrix}$$

Using Cramer's rule:

$$\begin{bmatrix} \frac{\partial \ln w_s}{\partial A_s} \\ \frac{\partial \ln w_u}{\partial A_s} \end{bmatrix} = \frac{1}{\tilde{\theta}_{S^h_m} - \tilde{\theta}_{S^\ell_m}} \begin{bmatrix} (\tilde{\theta}_{S^h_m} - \tilde{\theta}_{S^\ell_m})\frac{\partial \ln K^h_t}{\partial A_s} + (1-\tilde{\theta}_{S^h_m})(\frac{\partial \ln K_t}{\partial A_s} - \frac{\partial \ln K^h_t}{\partial A_s}) \\ (\tilde{\theta}_{S^h_m} - \tilde{\theta}_{S^\ell_m})\frac{\partial \ln K^h_t}{\partial A_s} - \tilde{\theta}_{S^\ell_m}(\frac{\partial \ln K^h_t}{\partial A_s} - \frac{\partial \ln K^h_t}{\partial A_s}) \end{bmatrix}$$

Hence:

$$\begin{bmatrix} \frac{\partial \ln w_s}{\partial A_s} \\ \frac{\partial \ln w_u}{\partial A_s} \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln K^h_t}{\partial A_s} \\ \frac{\partial \ln K^h_t}{\partial A_s} \end{bmatrix}$$

This equation means that skilled-biased factor-augmenting technical change in agriculture will result in wage increases for high and low skilled workers of the exact same

[37] Note that an alternative is to use the fact that the cost function is Cobb-Douglas as we have in the main text.

magnitude. Note that this result is a consequence of the small open economy assumption. If increased exports of low-skill intensive goods decreased prices of low-skilled intensive goods, then Stolper-Samuelson type forces would appear, which would tend to decreased low-skilled workers' wages.

We now turn to land prices. From,

$$0 = \theta_{T_a} \frac{\partial \ln w_T}{\partial A_s} + \theta_{S_a} \frac{\partial \ln w_s}{\partial A_s} - \theta_{S_a} \frac{\partial \ln A_s}{\partial A_s} + \theta_{U_a} \frac{\partial \ln w_u}{\partial A_s} - \frac{\partial \ln K_t^h}{\partial A_s}$$

we have that:

$$\frac{\partial \ln w_T}{\partial A_s} = \frac{(1 - \theta_{S_a} - \theta_{U_a})}{\theta_{T_a}} \frac{\partial \ln K_t^h}{\partial A_s} + \frac{\theta_{S_a}}{A_s \theta_{T_a}} = \frac{\partial \ln K_t^h}{\partial A_s} + \frac{\theta_{S_a}}{A_s \theta_{T_a}}$$

□

Proposition 3. *An increase in skilled-biased-factor-augmenting technology in agriculture (A_s), leads to an expansion of low-skill intensive manufacturing industries, provided that:*

1. *High- and low-skilled workers are imperfect substitutes (i.e. when $\varepsilon > 1$)*
2. *Land and labor are strong complements (i.e. when $\sigma < \varepsilon\Gamma$)*
3. *Agriculture is not much more intensive in low-skilled labor than the low-skill intensive industry.*

Proof. Consider the factor market clearing equilibrium conditions,

$$a_{Ta} Y_a = T \tag{24}$$

$$a_{Sa} Y_a + a_{S_m^\ell} Y_m^\ell + a_{S_m^h} Y_m^h = S \tag{25}$$

$$a_{Ua} Y_a + a_{U_m^\ell} Y_m^\ell + a_{U_m^h} Y_m^h = U \tag{26}$$

Log-differentiating Equations 24, 25 and 26 we get that:

$$a_{Ta} dY_a + da_{Ta} Y_a = dT$$

$$a_{Sa} dY_a + da_{Sa} Y_a + a_{S_m^\ell} dY_m^\ell + a_{S_m^h} dY_m^h = dS$$

$$da_{Ua} Y_a + a_{Ua} dY_a + a_{U_m^\ell} dY_m^\ell + a_{U_m^h} dY_m^h = dU$$

Now, define a hat-variable as $\widehat{X} = \frac{dX}{X}$ and $\lambda_{ij} = \frac{a_{Ij} Y_j}{I}$, i.e the share of factor I in industry j. Therefore, dividing at both sides of the equalities by the respective factor endowment, we can write the previous expressions as follows:

$$\lambda_{Ta}\widehat{Y}_a + \lambda_{Ta}\widehat{a}_{Ta} = \widehat{T} \tag{27}$$

$$\lambda_{Sa}\widehat{Y}_a + da_{Sa}\frac{Y_a}{S} + \lambda_{S_m^\ell}\widehat{Y_m^\ell} + \lambda_{S_m^h}\widehat{Y_m^h} = \widehat{S} \tag{28}$$

$$\lambda_{Ua}\widehat{Y}_a + da_{Ua}\frac{Y_a}{U} + \lambda_{U_m^\ell}\widehat{Y_m^\ell} + \lambda_{U_m^h}\widehat{Y_m^h} = \widehat{U} \tag{29}$$

Since in our economy the factor endowments are unchanged, $dT = dS = dU = 0$. This simplifies the expressions above in the following way:

$$\widehat{Y}_a = -\widehat{a_{Ta}} \tag{30}$$

$$\lambda_{Sa}\widehat{Y}_a + \lambda_{S_m^\ell}\widehat{Y_m^\ell} + \lambda_{S_m^h}\widehat{Y_m^h} = -da_{Sa}\frac{Y_a}{S} \tag{31}$$

$$\lambda_{Ua}\widehat{Y}_a + \lambda_{U_m^\ell}\widehat{Y_m^\ell} + \lambda_{U_m^h}\widehat{Y_m^h} = -da_{Ua}\frac{Y_a}{U} \tag{32}$$

Combining these expressions, we arrive to:

$$\lambda_{S_m^\ell}\widehat{Y_m^\ell} + \lambda_{S_m^h}\widehat{Y_m^h} = -\widehat{a_{Sa}}\lambda_{Sa} + \lambda_{Sa}\widehat{a_{Ta}} = \underbrace{\lambda_{Sa}(\widehat{a_{Ta}} - \widehat{a_{Sa}})}_{\gamma_s} \tag{33}$$

$$\lambda_{U_m^\ell}\widehat{Y_m^\ell} + \lambda_{U_m^h}\widehat{Y_m^h} = -\widehat{a_{Ua}}\lambda_{Ua} + \lambda_{Ua}\widehat{a_{Ta}} = \underbrace{\lambda_{Ua}(\widehat{a_{Ta}} - \widehat{a_{Ua}})}_{\gamma_u} \tag{34}$$

$$\widehat{Y_m^h} = \frac{\lambda_{U_m^\ell}\gamma_s - \lambda_{S_m^\ell}\gamma_u}{\Delta} \tag{35}$$

$$\widehat{Y_m^\ell} = \frac{\lambda_{S_m^h}\gamma_u - \lambda_{U_m^h}\gamma_s}{\Delta} \tag{36}$$

where $\Delta \equiv \lambda_{U_m^\ell}\lambda_{S_m^h} - \lambda_{U_m^h}\lambda_{S_m^\ell}$ and $\Delta > 0$ since the share of unskilled in the low-skilled intensive industry times the share of skilled in the skill-intensive industry is greater than the share of high-skilled in the low-skilled intensive industry times the share of unskilled in the high-skilled intensive industry. Then, $\widehat{Y_m^h} < 0$ iff $\lambda_{U_m^\ell}\gamma_s - \lambda_{S_m^\ell}\gamma_u < 0$. Which holds iff:

$$\lambda_{U_m^\ell}\gamma_s < \lambda_{S_m^\ell}\gamma_u$$

This can be re-written as:

$$\lambda_{U_m^\ell}\lambda_{Sa}(\widehat{a_{Ta}} - \widehat{a_{Sa}}) < \lambda_{S_m^\ell}\lambda_{Ua}(\widehat{a_{Ta}} - \widehat{a_{Ua}})$$

This can be further simplified to:

$$\lambda_{U_m^\ell}\lambda_{Sa}(\widehat{a_{Sa}}+\widehat{Y_a}) > \lambda_{S_m^\ell}\lambda_{Ua}(\widehat{a_{Ua}}+\widehat{Y_a})$$

And so, $\widehat{Y_m^h}<0$ iff:

$$\frac{\lambda_{U_m^\ell}}{\lambda_{S_m^\ell}}\frac{(\widehat{a_{Sa}}+\widehat{Y_a})}{(\widehat{a_{Ua}}+\widehat{Y_a})} > \frac{\lambda_{Ua}}{\lambda_{Sa}}$$

Now, note that $\widehat{a_{Sa}} > \widehat{a_{Ua}}$, which we show that it holds in more detail below (note, however, that this is simply saying that the demand for high-skilled labor increases relative to unskilled labor with increases in A_s). From this, we have that, $a^* \equiv \frac{(\widehat{a_{Sa}}+\widehat{Y_a})}{(\widehat{a_{Ua}}+\widehat{Y_a})} > 1$. Hence, we have that $\widehat{Y_m^h}<0$ iff $\frac{\lambda_{U_m^\ell}}{\lambda_{S_m^\ell}}a^* > \frac{\lambda_{Ua}}{\lambda_{Sa}}$. This condition holds as long as agriculture is not much more intensive in low-skilled labor than the low-skilled intensive industry.

Finally we are going to prove that $\widehat{a_{Sa}} > \widehat{a_{Ua}}$. This condition basically says that the elasticity of the agricultural unit factor demand with respect to A_s is larger for the skilled factor than for the unskilled factor, i.e $\frac{\partial ln a_{Sa}}{\partial ln A_s} > \frac{\partial ln a_{Ua}}{\partial ln A_s}$. Now, take the marginal productivities for skilled and unskilled labor in agriculture (Equations 21 and 22) and equate them to their factor price:

$$\begin{aligned} w_u &= MPU_a \\ w_s &= MPS_a \end{aligned}$$

and notice that we can write the following conditional labor demand equations:

$$\begin{aligned} U_a^{\frac{1}{\varepsilon}} &= \frac{1}{w_u}A_nK\gamma\Theta^{\frac{1}{\sigma-1}}A_L^{\frac{\sigma-1}{\sigma}}L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}}A_u^{\frac{\varepsilon-1}{\varepsilon}} \\ S_a^{\frac{1}{\varepsilon}} &= \frac{1}{w_u}A_nK\gamma\Theta^{\frac{1}{\sigma-1}}A_L^{\frac{\sigma-1}{\sigma}}L_a^{\frac{-(\varepsilon-\sigma)}{\varepsilon\sigma}}A_s^{\frac{\varepsilon-1}{\varepsilon}} \end{aligned}$$

Log-differentiating both expressions with respect to A_s :

$$\frac{\partial lnU_a}{\partial lnA_s} = \varepsilon\left[\frac{1}{\sigma-1}\frac{\partial ln\Theta}{\partial lnA_s} - \frac{(\varepsilon-\sigma)}{\varepsilon\sigma}\frac{\partial lnL_a}{\partial lnA_s}\right]$$
$$\frac{\partial lnS_a}{\partial lnA_s} = \varepsilon\left[\frac{1}{\sigma-1}\frac{\partial ln\Theta}{\partial lnA_s} - \frac{(\varepsilon-\sigma)}{\varepsilon\sigma}\frac{\partial lnL_a}{\partial lnA_s} + \frac{\varepsilon-1}{\varepsilon}\right]$$

Therefore,

$$\widehat{a_{Sa}} > \widehat{a_{Ua}} \iff \frac{\partial ln a_{Sa}}{\partial ln A_s} > \frac{\partial ln a_{Ua}}{\partial ln A_s} \iff \frac{\partial ln S_a}{\partial ln A_s} > \frac{\partial ln U_a}{\partial ln A_s} \iff \varepsilon - 1 > 0 \tag{37}$$

Therefore, $\widehat{Y_m^h} < 0$ and $\widehat{Y_m^\ell} > 0$. Upon the technical change in agriculture, the low-skill intensive industry expands and the high-skill intensive industry contracts.

□

Proposition 4. *When the following conditions hold:*

1. *High- and low-skilled workers are imperfect substitutes (i.e. when $\varepsilon > 1$)*
2. *Land and labor are strong complements (i.e. when $\sigma < \varepsilon\Gamma$)*
3. *Agriculture is not much more intensive in low-skilled labor than the low-skill intensive industry.*

An exogenous change in skill-biased-factor-augmenting technology (A_s), results in:

1. *Static gains from increased productivity in the agricultural sector.*
2. *Dynamic losses shaped by the decrease in the incentives to invest in new intermediate varieties for the H-industry.*

In particular, the growth rate of consumption is given by:

$$g_C = \frac{\max\{\pi^l, \pi^h, r\} - \rho}{\eta} \tag{38}$$

The change in gross domestic output is given by:

$$\frac{\partial \ln GDP_t}{\partial A_s} = \underbrace{\omega_a \frac{\partial \ln p_a Q_a}{\partial A_s} + \omega_m^\ell \frac{\partial \ln p_m^\ell Q_\ell}{\partial A_s} + \omega_m^h \frac{\partial \ln Q_h}{\partial A_s}}_{\textit{Static gains/losses}} + \underbrace{\mathbb{1}_{\{\pi^h \geq \pi^l\}} \frac{\chi}{\eta} \frac{\partial F_m^h}{\partial A_s} t}_{\textit{Dynamic gains/losses}} \tag{39}$$

where $\omega_j = \frac{p_j Q_j}{p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h}$.

Proof. Note that entrepreneurs produce intermediates for either the H- or the L-industries if the returns to entering are at least r. Hence, the growth rate of consumption is the maximum between the three possible investments.

$$g^C = \frac{\max\{\pi^\ell, \pi^h, r\} - \rho}{\eta}$$

In general, this is pinned-down by the profits made in the H-industry, hence in steady state:

$$g^C = \frac{\chi F_m^h(U_m^h, S_m^h) - \rho}{\eta}$$

This equation shows that consumption is growing in steady-state as a function of the size of the high-skilled sector. Moreover, knowledge grows at the level of investment, which is given by what is not consumed. The growth rate in each sector is given by the growth rate in K_t^h which is given by investment. This means that everything is growing at the same rate as consumption.

Finally we need to see how skilled-biased-factor-augmenting productivity increases affect the growth rate of the economy. For this, we obtain the evolution of GDP:

$$GDP_t = p_a K_t^h Q_a + p_m^\ell \varsigma K_t^h Q_m^\ell + \varsigma K_t^h F_m^h$$

to obtain that:

$$\ln GDP_t = \ln K_t^h + \ln(p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h)$$

In equilibrium, we have that $\ln K_t = \ln K_0 + g_c t$. And, hence:

$$\ln GDP_t = \ln K_0 + g^C t + \ln(p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h)$$

And hence:

$$\frac{\partial \ln GDP_t}{\partial A_s} = \frac{\partial g^C}{\partial A_s} t + \frac{\partial \ln(p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h)}{\partial A_s}$$

And hence

$$\frac{\partial \ln GDP_t}{\partial A_s} = \frac{\partial g^C}{\partial A_s} t + \frac{1}{p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h} \left(\frac{\partial Q_a}{\partial A_s} + \frac{\partial p_m^\ell \varsigma Q_m^\ell}{\partial A_s} + \frac{\partial \varsigma F_m^h}{\partial A_s} \right)$$

And hence:

$$\frac{\partial \ln GDP_t}{\partial A_s} = \frac{\partial g^C}{\partial A_s} t + \omega_a \frac{\partial \ln p_a Q_a}{\partial A_s} + \omega_m^\ell \frac{\partial \ln p_m^\ell Q_m^\ell}{\partial A_s} + \omega_m^h \frac{\partial \ln F_m^h}{\partial A_s}$$

with $\omega_a = \frac{p_a Q_a}{p_a Q_a + p_m^\ell \varsigma Q_m^\ell + \varsigma F_m^h}$, and analogously for the other ω_j.

Which is equal to:

$$\frac{\partial \ln GDP_t}{\partial A_s} = \underbrace{\omega_a \frac{\partial \ln p_a Q_a}{\partial A_s} + \omega_m^\ell \frac{\partial \ln p_m^\ell Q_\ell}{\partial A_s} + \omega_m^h \frac{\partial \ln Q_h}{\partial A_s}}_{\text{Static gains/losses}} + \underbrace{\mathbb{1}_{\{\pi^h \geq \pi^l\}} \frac{\chi}{\eta} \frac{\partial F_m^h}{\partial A_s} t}_{\text{Dynamic gains/losses}}$$

$\square$