The Agricultural Wage Gap: Evidence from Brazilian Micro-data*

Jorge A. Alvarez†

March 29, 2017

Abstract

A key feature of developing economies is that wages in the agricultural sector are significantly below those of other sectors. Using a panel data set on the universe of formal workers in Brazil, I use information on workers that switch sectors to decompose the drivers of this inter-sector gap. I find that most of the gap between sectors is explained by unobservable differences in the skill composition of workers, as opposed to differential pay of workers with similar skills. The evidence speaks against the existence of large short-term wage gains from the reallocation of workers out of agriculture and favors recently proposed Roy models of inter-sector sorting as drivers of lower average wages in agriculture. A calibrated model of worker sorting can account for the wage gap observed in 1996 Brazil and a share of both the wage gap decline and the diminishing worker participation in agriculture observed during the period between 1996 and 2013.

Key words: Wage Gaps, Productivity Gaps, Structural Transformation, Agriculture, Human Capital, Sorting, Brazil.

*I am grateful for the insightful and generous advice of Richard Rogerson throughout this project. I am also thankful to Stephen Redding, Elhanan Helpman, and Marc Muendler for granting me access to part of the data during the earlier stages of this project. In addition, I want to give special thanks to Niklas Engbom and Chris Moser for their collaboration in working on Brazilian labor markets. I also appreciate the input of Angus Deaton, Douglas Gollin, Ilyana Kuziemko, Oleg Itskhoki, David Lagakos, Prakash Loungani, Edward Miguel, Ben Moll, Chris Papageorgiou, Stephen Redding, Tom Vogl, as well as seminar participants of the Princeton Macroeconomics lunch seminar, the Princeton Research Program for Development Studies and Center for Health and Wellbeing lunch seminar, the Princeton Public Finance Working Group and the International Monetary Fund's Development Macroeconomics seminar for their helpful comments on my work on both agriculture and Brazilian wage differences. The views expressed in this study are the sole responsibility of the author and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

†Mail: International Monetary Fund, 700 19th Street NW, Washington, DC 20431. Email: jalvarez@imf.org.

1 Introduction

A key feature of developing economies is that wages in the agricultural sector are significantly below those of other sectors.[1] Additionally, these economies have most of their workforce in the agricultural sector. These two observations motivate a literature dating back to Lewis (1955) and Rostow (1960) that views the exit of workers out of agriculture as a fundamental mechanism of development. The body of work on agricultural development and inter-sector differences, however, has not completely settled the question of why so many workers stay in agriculture in spite of better wages being paid in other sectors. One possibility is that some barrier prevents the movement of workers across sectors, in which case wage gaps between agriculture and other sectors indicate unexploited potential gains from the reallocation of workers out of agriculture. A second possibility is that workers in agriculture are characteristically different from those in non-agriculture, in which case wage gaps would not be evidence of potential wage gains. The objective of this paper is to shed light on which of these possibilities is a more likely explanation of the agricultural wage gap.

A challenge in exploring this question is assessing the role of unobserved worker characteristics. For instance, if an agricultural worker and a non-agricultural worker with the same observable characteristics (e.g. age and education) earn different wages, it is hard to distinguish whether the two sectors have differential pay for similar workers or whether the two workers are different due to unobserved characteristics. This paper assesses the role of unobserved characteristics by using panel micro-data covering all sectors of the Brazilian economy from 1996 to 2013. The use of panel data is an improvement on the literature on agricultural wage gaps in developing countries, which has typically relied either on the estimation of structural models to match country-level moments or on the analysis of heterogeneous cross-sectional surveys from a sample of countries. Specifically, the panel dimension of the data allows me to control for differences in both observable and fixed unobservable worker characteristics. Information on workers that switch between sectors (from now on referred to as 'sector-switchers') can be used to distinguish whether the wage gap between agriculture and non-agriculture reflects differential pay of similar workers in the two sectors or, alternatively, whether the gap is due to differences in the composition of worker characteristics in each sector.

The main empirical finding of this study is that workers who transition out of agriculture

[1] In a sample of developing countries studied by Vollrath (2014), the median average wage ratio between agriculture and manufacturing was 1.6. This is 1.9 when comparing agriculture against services. In the sample of countries studied by Herrendorf and Schoellman (2015), the median ratio between agriculture and the rest of the economy is 2.0.

experience limited compensation gains when compared to the overall gap in mean wages between agriculture and other sectors. I conclude that the agricultural wage gap does not appear to be driven by differential pay of similar workers, once fixed unobservable characteristics are controlled for. Instead, the largest share of the agricultural wage gap is explained by differences in the composition of worker characteristics in each sector. In addition, I find that the wage gap between agriculture and other sectors in Brazil declined significantly from 1996 to 2013 as the economy grew richer. This reduction is similar when comparing agriculture to both services and manufacturing, and it coincided with a decline in the share of workers employed in agriculture — from 25 percent to 14 percent. Moreover, this decline does not appear to be driven by changes in educational attainment or country demographics. In fact, I find that age and education explain only a small share of the large wage gap in Brazil during the late 1990s, and that differences in the composition of these variables between sectors drove only a small share of the decline during this period. Most of the decline is driven by compositional changes in the distribution of fixed unobservable worker characteristics.

Both the limited wage gains from transitions out of agriculture and the importance of worker composition differences between sectors pose a challenge for an agricultural wage gap model. Such a model must generate large declining wage gaps that do not result in large wage gains among sector-switchers. Building on the work of Roy (1951), a recent literature has proposed worker sorting as a possible explanation that is consistent with this pattern. In particular, Lagakos and Waugh (2013) and Young (2013) illustrate how workers with sector-specific skills can sort themselves into different sectors to generate large wage gaps. In this type of model, each worker faces a choice between two idiosyncratic wages in agriculture and non-agriculture. Workers with a comparative advantage in non-agriculture choose to work in that sector, and this generates a wage gap relative to workers who find it advantageous to stay in the agricultural sector.

To test the explanatory power of this mechanism, I build on the sorting model proposed by Lagakos and Waugh (2013) and test whether a calibrated model that targets micro-moments from sector-switchers can generate wage gaps of the magnitudes observed in Brazil in 1996. I find that a large wage the gap level can be generated by this model. In a second stage of analysis, I use the model to explore productivity growth, technological change and a compression of skill differences as potential drivers of the wage gap decline. I find that the latter two can generate a qualitative decline, though none of these factors generate a fall of the magnitude observed in Brazil during the period between 1996 and 2013.

The rest of the paper is structured as follows. Section 2 provides a literature review that relates this paper to the labor literature on inter-sector wage gaps and the macroeconomic literature on both wage and output per worker gaps between sectors. Section 3 describes the datasets used. Section 4 describes the magnitude and evolution of the wage and productivity gaps in Brazil as well as the decline in the share of workers employed in the agricultural sector. Section 5 assesses the role of observables, unobservables, and differential pay of similar workers in explaining the gap. Section 6 describes the mechanics and calibration of an economy where workers sort across sectors, as well as the power of worker sorting in explaining the agricultural wage gap magnitude and its decline. Section 7 concludes.

2 Literature Review

Most studies show that large inter-sectoral wage gaps persist even after controlling for educational attainment and other worker observables. The remaining gap stems from either differential pay of similar workers or, alternatively, sector differences in the composition of both observable and unobservable worker characteristics.

U.S. labor studies have explored this distinction with mixed results. Using matched data from the Consumer Population Survey (CPS), Krueger and Summers (1988) argue that unobservable worker characteristics cannot explain much of the difference in wages between sectors. On the other hand, Murphy and Topel (1987; 1990) also use the CPS and conclude that industry switchers receive only 27 to 36 percent of the total industry differential, and thus nearly two-thirds of inter-sector wage gaps can be attributed to differences in the composition of worker characteristics in each sector. Also using US data, Gibbons and Katz (1992) find limited evidence for differential pay of similarly-skilled workers between sectors and instead highlight the role of differences in the composition of observable and unobservable characteristics.

International studies on developing countries have also highlighted the role of differences in observable and unobservable worker characteristics in explaining the gap. Vollrath (2014) finds that large wage differences exist between workers after controlling for observed human capital in a set of 14 countries. He explores whether these wage gaps could be the result of distortions that prevent workers from being paid the value of their marginal product in each sector. Using a misallocation framework similar to Hsieh and Klenow (2009), Vollrath (2014) estimates that potential gains from eliminating distortions and eradicating human capital misallocation are less than five percent in developing countries. If misallocation is not important, this implies that differences in the composition of worker productivity are likely to be more important drivers of the gap. Similarly, using a different sample of countries, Herrendorf and Schoellman (2015) regress wages on observables allowing for returns on observables to vary by sector. They conclude that most of the wage gap between agriculture and other sectors can be accounted for by differences in workers' human capital — and sector-specific differential returns— present in each sector.

However, because of data constraints, these studies are limited to the comparison of a diverse collection of cross-sectional surveys. This prevents rigorous empirical testing of whether differences attributed to unobservable characteristics or differential human capital returns could in fact be the result of other forces producing differential pay of similar workers. Mobility frictions and compensating differentials, for instance, are two alternative explanations consistent with both the differential returns on observables estimated by Herrendorf and Schoellman (2015) and the residual wage differences reported by Vollrath (2014). By using a panel dataset where workers are observed as they switch across sectors, the current study overcomes the limitations of cross-sectional data and distinguishes the role of fixed unobservable characteristics from alternative stories of differential pay. This approach has been recently used by Hendricks and Schoellman (2017) to study gains from migrations[2] and by Hicks et al. (2017) to study sectoral wage gaps using panel data from Indonesia and Kenya. Consistent with this paper, they find limited gains from sectoral transitions when compared to larger aggregate wage gaps.

The study of wage gaps is also closely related to the study of output per worker gaps between agriculture and other sectors. Kuznets (1971), Caselli (2005), Restuccia, Yang and Zhu (2008), among others, have argued that a large share of income differences across countries is explained by labor productivity gaps between agriculture and other sectors. However, focusing on output per worker, even in advanced countries, risks exposure to important sources of measurement errors. For instance, Gollin, Parente, and Rogerson (2004) suggest that unaccounted home production understates agricultural output and Herrendorf and Schoellman (2015) point out that errors in value added measurement muddy comparisons of worker productivity across US states. Partially as a result of this, the role that both observed and unobserved human capital play in explaining these output per worker gaps is still an open debate. Herrendorf and Schoellman (2015) argue that human capital accounts for most of the output per worker gap between agriculture and other sectors in the US and other selected countries. Gollin, Lagakos, and Waugh (2014) argue that human capital —along with adjustments to labor supply— account for only about a third of the gap in the developing countries they study. Focusing on wages avoids many of the problems

[2]Other studies on migration include Beegle, Weerdt and Dercon (2011), Bryan, Chowdhury and Mobarak (2014), Chiquiar and Hanson (2005), and Yang (2006).

with the measurement of differences between agriculture and the rest of the economy. Although wages and output per worker are not equivalent measures of labor productivity, the results of this paper can speak to some of the debates about the role of differences in worker composition on inter-sector gaps explored by this literature.

Beyond establishing the role of worker characteristics in explaining the inter-sector gaps, a second objective of the literature is to uncover the mechanisms behind compensation and output per worker differences. Two main types of mechanisms are relevant to this study. The first are distortions that create wedges in marginal productivity of labor between sectors. These distortions can include scale effects that impact the allocation of resources across agricultural firms (Adamopoulos and Restuccia (2014), Donovan (2016)) or barriers that prevent the free flow of capital and workers (Restuccia and Rogerson (2008*a*), Herrendorf and Teixeira (2011)). Distortions that prevent marginal labor products to equalize have also been studied at the firm level by Restuccia and Rogerson (2008*b*) and Hsieh and Klenow (2009), who highlight their greater importance in developing countries. To the extent that these distortions are also present between sectors —and workers are not freely mobile— the mechanisms generating productivity gaps can be related to the agricultural wage gap.

A second type of mechanism highlighted by Young (2013), and Lagakos and Waugh (2013) portrays wage gaps as the result of sector differences in worker skill composition. Lagakos and Waugh (2013) illustrate how such skill differences can be the result of an equilibrium outcome. In their model, workers sort themselves to the sector where they are most productive. This process induces differences in the composition of worker skills employed by each sector, and this in turn generates a gap in mean wages paid in agriculture relative to non-agriculture. Importantly, the agricultural gap in this context is not the result of any additional distortions that induces differential pay of similar workers. Building on this idea, Young (2013) uses cross-sectional surveys from developing countries to show how migration is consistent with rural-urban consumption driven by the sorting of workers. Although his focus is on consumption, his findings are also consistent with agricultural wage gaps generated by the sorting of workers with different unobservable skills. The mechanism proposed by this paper —which is also supported by the empirical results to be presented— belongs to this family of sorting models, where the agricultural wage gap is ultimately driven by compositional differences in worker characteristics.

3 Data description

Two main databases are used. The first is the set of Brazilian household surveys from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD) from 1996 to 2013. This contains a representative sample of households covering all of Brazil. The survey includes both formal and informal workers and records demographic and employment-status characteristics as well as monthly earnings for all members of a household. In this paper, this data is used to show trends in earnings among all workers, including both formal and informal, during the period of study.[3] In particular, I establish that the trends and magnitudes in inter-sector pay differences in Brazil among all workers are similar to the ones observed among formal workers. Data from PNAD is also used to compute the total number of workers in each sector and —in combination with the national accounts recorded by the *Instituto Brasileiro de Geografía e Estadística* (IBGE)— value added per worker for each year and sector. Due to the cross-sectional nature of the surveys, however, individuals cannot be followed over time in the PNAD. I am therefore unable to control for worker unobservable characteristics using data on both formal and informal workers. For this reason, most empirical decompositions in this paper focus on formal sector data which is now described in greater detail.

Data on formal workers comes from the *Relação Anual de Informações Sociais* (RAIS), which is administered by the Brazilian Ministry of Labor and Employment. This database is constructed from a mandatory annual survey filed by all formally registered firms in Brazil and contains earnings, occupation and demographic characteristics of workers as reported annually by their employers.[4] Importantly, each worker in the data has a unique and time-invariant worker ID that does not change as workers switch employers. This feature of the data allows me to follow individuals over time and create a panel of the universe of employed formal workers across all sectors. In addition, each worker is linked to their employing firm, which also has a unique and time-invariant ID. This allows me to link workers to their respective sectors, and identify transitions between sectors.[5] The data covers the period from 1996 to 2013.[6]

The RAIS dataset reports average monthly gross labor earnings including regular salary

[3]Because hours data is only reported in broad categories in the PNAD, I focus on earnings when comparing trends among formal and informal workers in the economy.

[4]It is common practice for businesses to hire a specialized accountant to help with the completion of the RAIS survey to avoid fines levied on late, incomplete, or inaccurate reports, which makes the quality of the data superior to household surveys.

[5]IDs available are anonymized to protect the identity of both workers and firms.

[6]Although earlier years are available for a large subset of Brazilian workers, the lack of universal coverage in earlier periods can be particularly problematic in studying transitions out of agriculture. Hence, the analysis is restricted to this later period.

payments, holiday bonuses, performance-based and commission bonuses, tips, and profit-sharing agreements as well as the start and end month of the job. To account for heterogeneity in the duration of job-spells, I divide annual earnings by the number of months worked at each job within a particular firm to get a measure of monthly earnings. This is divided by hours contracted per month to get a measure of hourly wages. A worker might have multiple spells in a year if he or she switched employers during the year or worked multiple jobs, but on-the-job earnings changes within a year are not recorded. To standardize the dataset at an annual level, I restrict attention to a unique observation per worker-year by choosing the highest-paying among all employment spells in any given year.

The dataset also contains the age and educational attainment of each worker. Educational levels are classified into less than high school, high school, some college education, and completed college education. In all regression specifications utilizing age and education as explanatory variables of the wage gap, a full set of age and education interacted dummies is used.

Finally, to identify the employment sector and occupation of workers, classification is based on categories from the IBGE. Both the industry and occupation classification system changed during the period of study. Here, I use conversion tables provided by IBGE to standardize classification between different years and choose categories for both occupations and sectors coarse enough in order to avoid potential biases arising from mechanical changes in the classification system over time. The three sectors used are Agriculture, Manufacturing (including energy and mining), and Services. Occupation categories used are at the three-digit disaggregation level.

Due to imperfect matching of all categories within a sector and occupation classification system, I exclude firms with inconsistent sector classifications so that sector switchers are not incorrectly specified. I also exclude individual observations that have either firm IDs or worker IDs reported as invalid as well as data points with missing wages, dates of employment, educational attainment, hours, or age. For computational purposes, a ten percent sample is used in all estimations. This includes more than three million workers and more than ten thousand sector-switchers in any given year.

For all estimations, I restrict the analysis to workers between 18 and 65 years old with contracted hours of at least 30 hours a week. Table 1 provides key summary statistics for the RAIS data for three sub-periods: 1996-2001, 2002-2007, and 2008-2013. Some features of the data are worth noting. The first is that the number of workers increases substantially over time from 4.8 million workers in 1996-2001 to 7.8 million in 2008-2013. This rise is mainly the result of two forces: population growth and an increase in formality in Brazil. A second observation is that education is quite different in agriculture in Brazil relative to other sectors. In 1996-2001, for instance, only five percent of formal workers in the agricultural sector had a high school degree and one percent had completed college, relative to 34 and ten percent in other sectors. During 1996-2013, educational attainment substantially improved partially as a result of educational reforms in the late 1990s and the rise of social programs in the 2000s. In contrast, the age distribution in each sector did not change substantially. The explanatory power of age and education will be one of the focal points of the analysis. Finally, though wages between agriculture and other sectors are quite different, there are only small gaps in earnings and wages when comparing services against manufacturing in all periods. This motivates the dual economy focus of this paper: explaining the gaps between agriculture and all other sectors in the economy.

4 The magnitude and evolution of the agricultural gap in Brazil

Differences in pay between agriculture and other sectors are large in Brazil, and these were significantly reduced during the last two decades. The ratio of mean earnings between non-agriculture and agriculture among all workers (both formal and informal) in the economy —as measured by the PNAD household surveys— declined from 2.2 in 1996 to 1.7 by 2013[7]. As discussed above, the main contributions of this paper hinge on the use of the panel structure of the data so that workers can be followed over time. Since this feature is only available for formal workers, the rest of the paper will focus on formal sector data. Similarly to the overall economy, formal workers exhibit a very similar decline in the ratio of mean earnings between non-agriculture and agriculture from 2.3 in 1996 to 1.6 in 2013 (Figure 1). The corresponding gap in hourly wages during the same period fell from 2.3 to 1.7. Moreover, the magnitude of the gap and its decline has been similar when comparing agriculture to both services and manufacturing individually. In contrast to the differences between agriculture and non-agriculture, mean earnings in the two non-agricultural sectors were similar throughout this period.

Another feature of the data is that the agricultural wage gap is present throughout the wage distribution. Figure 3 shows the ratio of wage percentiles in agriculture and non-agriculture. Percentiles are here defined by the ranking of workers within each sector.

[7] Earnings from PNAD surveys correspond to income from all jobs. Because of both the difficulty of hours measurement in the informal sector and the fact that PNAD only contains hours in broad categories, we use earnings time trends to establish trends in inter-sector gaps among formal and all workers in the economy.

Table 1: RAIS Summary Statistics

				log(Wages)		Education		Age	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	Sector	# Worker-years	# Unique Workers	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
1996-2001	Agriculture	0.97	0.31	1.25	0.61	1.10	0.42	34.36	11.21
	Manufacturing	3.53	0.96	1.87	0.84	1.43	0.74	32.33	9.86
	Services	12.62	3.56	1.90	0.86	1.69	0.93	34.41	10.67
	All	17.12	4.83	1.86	0.86	1.60	0.89	33.98	10.57
2002-2007	Agriculture	1.28	0.39	1.40	0.52	1.15	0.46	34.49	11.07
	Manufacturing	4.41	1.12	1.85	0.77	1.62	0.78	32.60	10.01
	Services	16.34	4.37	1.87	0.79	1.90	0.96	34.70	10.88
	All	22.03	5.89	1.84	0.78	1.80	0.92	34.27	10.76
2008-2013	Agriculture	1.34	0.38	1.69	0.49	1.28	0.57	35.70	11.22
	Manufacturing	5.53	1.39	2.04	0.71	1.81	0.80	33.34	10.37
	Services	22.76	6.06	2.04	0.73	2.04	0.94	35.04	11.18
	All	29.63	7.83	2.02	0.72	1.96	0.91	34.75	11.05

Note: 10 percent sample from all formal workers in Brazil.Number of workers and worker-years are in millions. Wages refer to average monthly earnings divided by hours in real terms (Using 2013 Reais). Education levels are defined as 1= Primary or middle school or no education, 2= high school 3= some college education and 4= college completed. Age is in years.

There is a pattern, with the top earners in the agricultural and non-agricultural sectors being further apart than the bottom earners in the two sectors. The differences, however, are still significant across all percentiles and it is not the case that wage gaps are a phenomenon that is only applicable to certain parts of the wage distribution. Furthermore, when looking at the evolution of these ratios over time, the decline in compensation differences does not appear to be driven by the catch up of only the poorest or richest parts of the distribution of agricultural workers.

In addition, the wage gap decline was accompanied by a similar decline in the value added per worker gap. Figure 2 shows how the between-sector difference in gross domestic product per worker as measured by the national accounts declines over the 1996-2013 period. Similar to the wage pattern, the decline is large when comparing agriculture against both manufacturing and services. Unlike wages, however, the differences and the magnitude of the decline is much larger when looking at the agriculture-manufacturing gap than when looking at the agriculture-services one. This is expected due in part to the natural differences in capital intensities between services and agriculture. These differences notwithstanding, the qualitative pattern of pay and value added per worker gaps is qualitatively similar. Importantly, this reduction in sectoral inequities occurred during a period where yearly real GDP growth averaged 2.7 percent, as the country transitioned out of a period of macroeconomic instability and hyperinflation into a period of technology modernization and growth.[8] The interrelation of growth, productivity, and the decline in inter-sector gaps will be central to our analysis of mechanisms in section 6.

The magnitudes of both the wage and value added per worker gaps between agriculture and other sectors are large when compared with other estimates in the literature. In 1996, the magnitude of the value added per worker gap between agriculture and other sectors is 5.3, which is greater than the maximum found by Herrendorf and Schoellman (2015) in their 12 country sample and just below the mean gap reported in Gollin, Lagakos and Waugh (2014) for the poorest quartile of countries in their 151 country sample. By 2013, after a cumulative real output growth of 61 percent, the value added per worker gap is 2.4. This estimate is similar to the median of 2.3 in the Herrendorf and Schoellman (2015) sample and closer to the 2.0 mean of the richest 25 percent of countries in the Gollin, Lagakos and Waugh (2014) sample. When compared to the cross-country evidence, Brazil appears to have endured a significant transformation during the period of study.

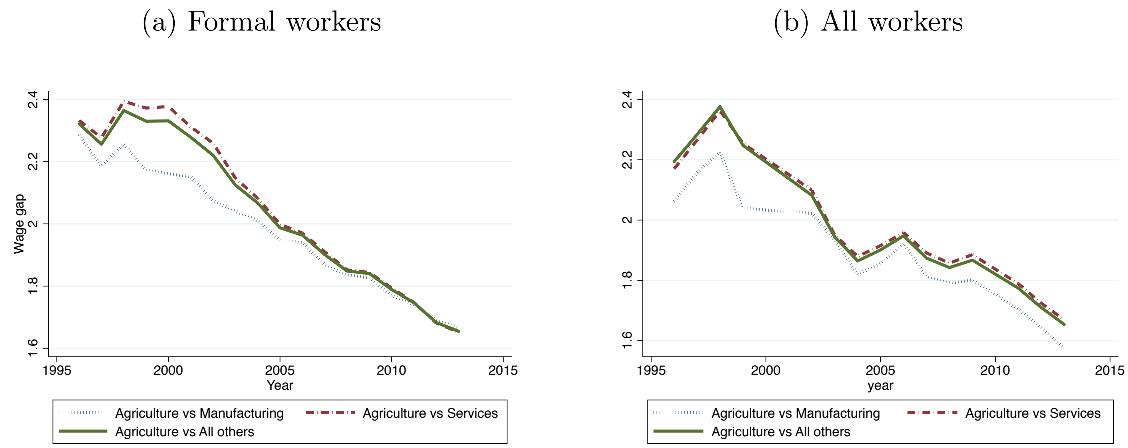
In terms of the wage gap between agriculture and other sectors, Brazil's 1996 wage gap

[8]Bustos, Caprettini and Ponticelli (2016) explain some of the agricultural modernization of the agricultural sector in Brazil.

of 2.3 is above the median of 2.0 from the Herrendorf and Schoellman (2015) sample. By 2013, it falls below the sample's mean to 1.6. Figure 4 shows how these gaps compare to the list of 15 developing countries studied by Vollrath (2014). Brazil's 1996 gap between agriculture and manufacturing would rank third highest, just below Ecuador in that sample. When comparing agriculture vs services, the rank would be 5th, just above Indonesia. In contrast, Brazil's 2013 gap levels with respect to manufacturing and services would rank 8th and 11th, respectively. Although the data on Brazil is not entirely comparable to the wage data from other countries' surveys, the significant move down the ranking of countries suggest that Brazil's decline cannot be described as an insignificant change.

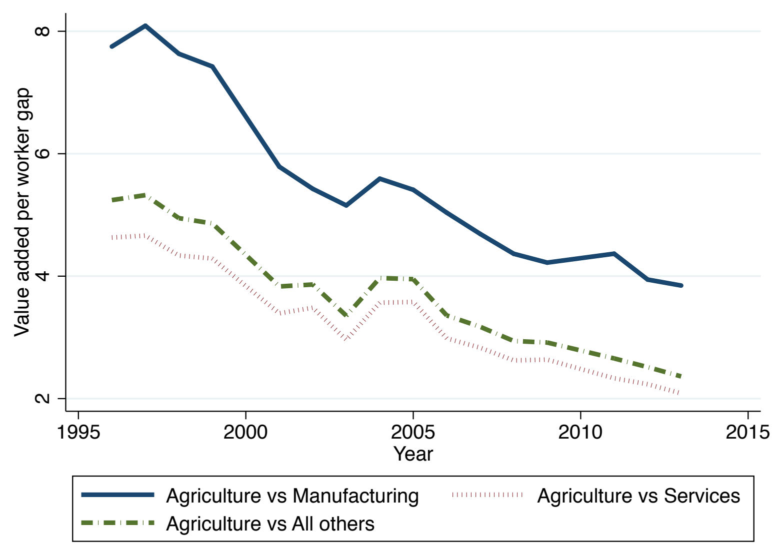
In parallel to the closing of both output per worker and wage gaps, Brazil also endured a substantial transformation of the employment structure. The workforce composition based on household surveys is shown in Figure 5. The economy employed 25 percent of the labor force in agriculture in 1996, which declined to 14 percent by 2013. Manufacturing employed 13 to 15 percent throughout this same period, and services increased from 61 to 72 percent. Among formal workers, a similar pattern is observed and the share of workers in agriculture has declined from 5.1 to 3.6 percent since 1996. Although the population of formal workers is much smaller than the universe of workers in agriculture, the magnitude of the wage gap also shows a declining pattern in the share of labor employed in agriculture. The interrelation between the movement of workers out of agriculture and the agricultural wage gap will be considered in section 6, when mechanisms behind the gap's decline are discussed. First, a statistical decomposition of the agricultural wage gap is conducted using the panel structure of the data on formal workers.

Figure 1: Wage gap in Brazil



Note: The wage gap is calculated as the ratio in average labor monthly earnings between agriculture, manufacturing and services as classified by the IBGE. Data on formal workers comes from the *Relação Anual de Informações (RAIS)*. Data on all workers (both formal and informal) comes from the PNAD household surveys.

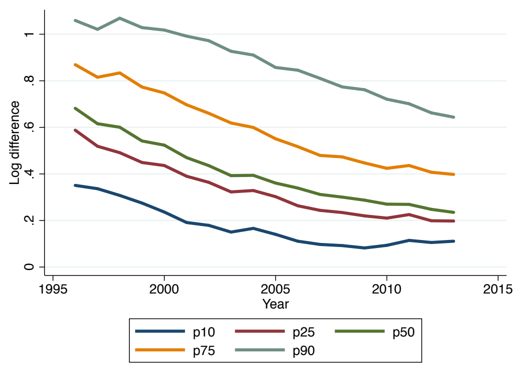
Figure 2: Value added per worker gap in Brazil



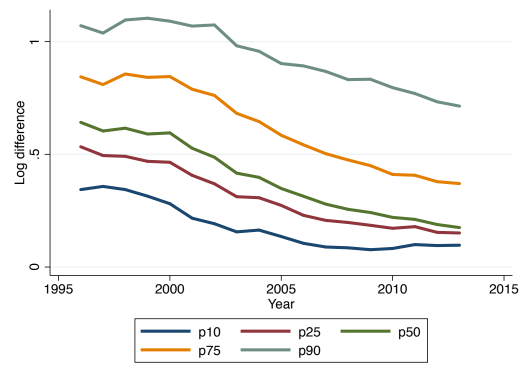
Note: Value added per worker gaps are constructed from national accounts available from IBGE and labor statistics from the Pesquisa Nacional por Amostra de Domicílios (PNAD).

Figure 3: Gaps in Brazil by percentile

(a) Agriculture vs manufacturing

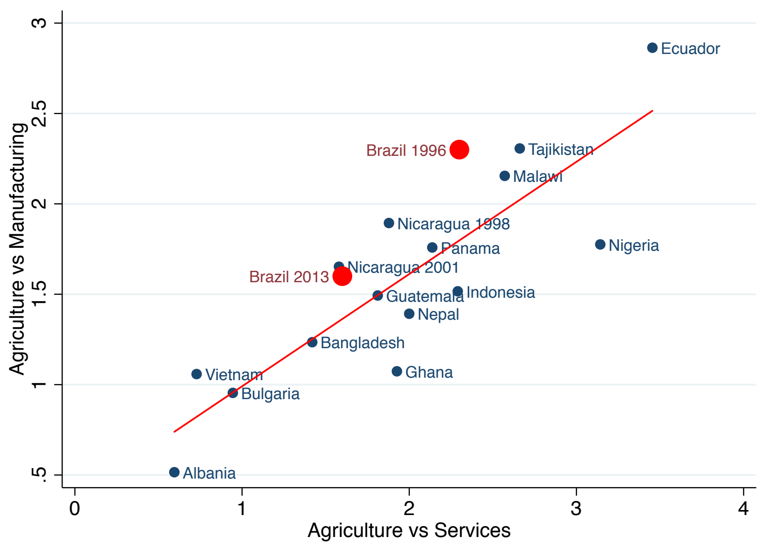


(b) Agriculture vs services



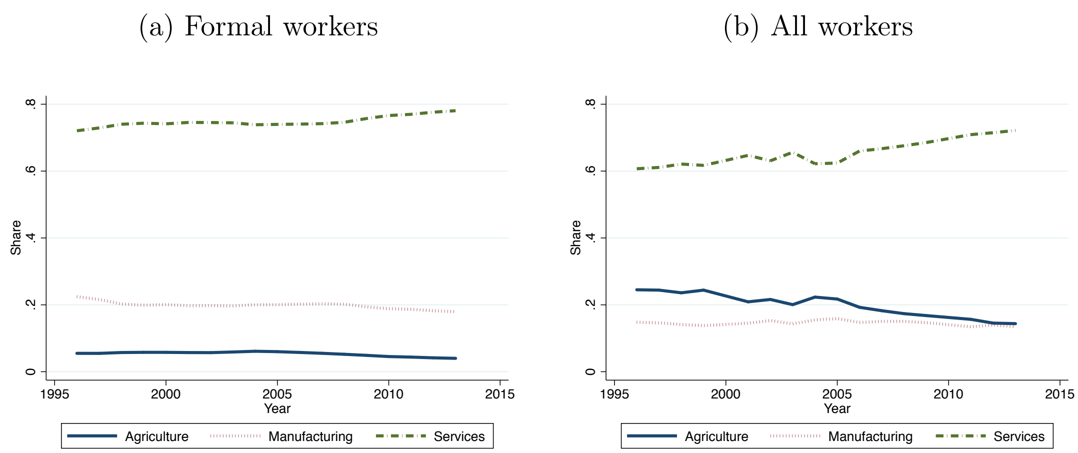
Note: Difference in the means of log wages between sectors for formal workers are presented. Each line corresponds to the difference between each percentile group in the two sectors.

Figure 4: Wage gaps in Brazil vs other countries



Note: Data for Brazil comes from PNAD and national accounts from IBGE. For other countries, estimates are constructed based on cross-country data from Vollrath (2014).

Figure 5: Workers by sector



Note: Share of total employed workers. Formal worker data is from RAIS. Data on all workers is from PNAD household surveys.

5 Sources of the agricultural gap

We now turn to explore what drives the wage gap between agriculture and other sectors. Three possible alternatives are considered. The first are differences in the composition of observable human capital as measured by age and education. The second are differences in the distribution of fixed unobserved worker characteristics between sectors. Finally, the third alternative is the presence of mechanisms that induce differential pay of similar workers employed by different sectors. Inter-sector mobility frictions, sector-specific rent-sharing agreements, and compensating differentials are some of the mechanisms that fit this third category. This section argues that the first two alternatives, where the gap is driven by compositional differences in worker characteristics, explain most of the agricultural wage gap and its decline.

5.1 Human capital

Differences in human capital introduce heterogeneity in the productivity of workers which, in a standard competitive environment, should translate into wage differences. Table 1 indeed shows differences in education between sectors, with agricultural workers being on average less educated than their peers in services and manufacturing. To the extent that these characteristics determine human capital, these differences can potentially explain part of the agricultural wage gap.

There are two margins on which human capital influences the wage gap. On the one hand, human capital can be lower in one sector than the other. On the other hand, even if the *composition* of human capital is the same in the two sectors, the *returns* to human capital might be different in the two sectors. I first assess whether compositional differences in human capital, as measured by age and education, can account for a substantial share of the gap by estimating the following model for each sector and year.

$$log(w_{ist}) = F_{st}(education_{ist}, age_{ist}) + \epsilon_{ist}$$

Here, $w_{ist}, education_{ist}, age_{ist}$ are the wage, education level, and age of worker i in sector s in year t. To impose minimal restrictions on how age and education influence wages, the mapping of education and age to wages is specified as $F_{st}(education_{ist}, age_{ist}) = \sum_{a,e} \mathbf{1}(age_{ist} = a, edu_{ist} = e) \times \beta^{s}_{aet}$. Thus, the specification allows full flexibility in terms of both age and education, and this relationship can vary in every year of the sample.

For the rest of the paper, I will define the wage gap as the mean difference of log hourly wages with respect to agriculture. Specifically, the gap between sector s' and agriculture is defined as

$$\Delta_{s'} E(log(w_{ist})) \equiv E(log(w_{ist})|s = s') - E(log(w_{ist})|s = a)$$

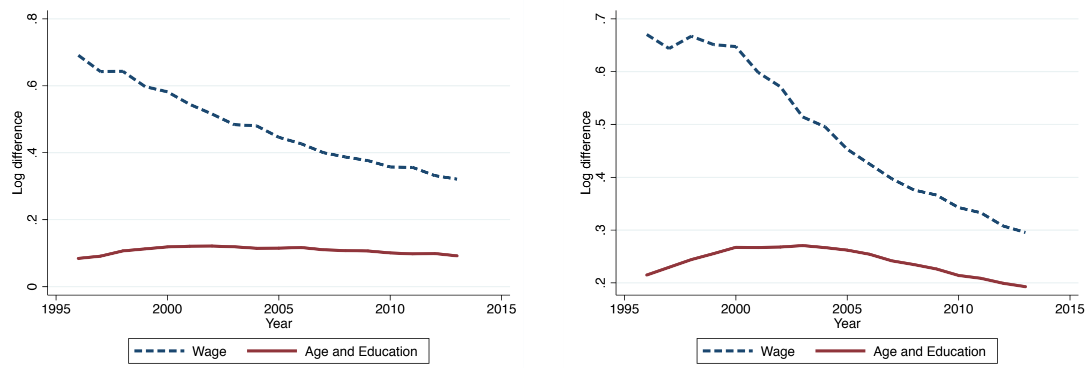
where the possible values for sector s', $\{a, m, s\}$, refer to agriculture, manufacturing and services respectively. The focus on additively separable mean log-wage gaps is used to simplify the presentation of the log-linear models to be studied.

Figure 6 shows the decomposition of the mean log difference into two parts: a component due to age and education and another due to the residual. There, we can see that the effect on wages from age and education differences between agriculture and other sectors have remained roughly constant throughout 1996-2013. When comparing agriculture and manufacturing, these observable characteristics explain a nearly constant 9–11 log points of the gap. When comparing agriculture and services, observables matter more and wage gaps due to age and education have averaged 24 log points. Overall, age and education differences accounted for ten to 26 percent of the wage gap level during the period. The results show that most of the wage gap level is largely driven by factors not accounted by compositional differences in age and education alone.

Moreover, the *decline* in the wage gap cannot be entirely attributed to changes in education and the distribution of age in each sector. When comparing manufacturing and agriculture, the stability of the gap due to age and education shown in Figure 6 contrasts the decline in the overall wage gap. When comparing services and agriculture, age and education explain some of the decline. However, the flatter pattern of this component relative to the total gap decline indicates that this reduction is not sufficient to explain the entire decline.

Figure 6: Gap in mean log wages between agriculture and other sectors due to age and education

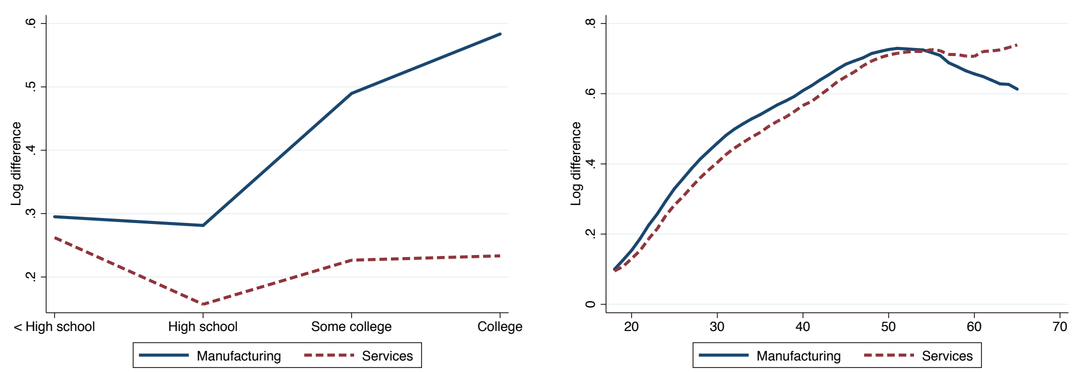
(a) Agriculture vs Manufacturing (b) Agriculture vs Services



Note: Wage refers to the difference in mean log wages between sectors. Age and education refer to the difference of the mean predicted values, $E(F_{st}(education_{ist}, age_{ist})|s = a) - E(F_{st}(education_{ist}, age_{ist})|s = a)$.

Figure 7: Mean difference in log wages relative to agriculture by educational attainment and age

(a) By education (b) By age



Note: Mean wage difference between manufacturing/services and agriculture by educational attainment and age.

The results above point to the importance of differences in pay within each education-age group across sectors. Figure 7 shows that average wage differences by education and age groups are large, with older workers gaining significantly less in agriculture relative to other sectors and workers in each age and education group being paid less than their comparable peers in non-agriculture. The difference in average pay for worker characteristics in each sector may reflect differential returns to education and experience by sector. For instance, worker with a high school degree might be more productive in manufacturing and services than in agriculture due to the availability of jobs that require this level of educational attainment.

The question is then to what extent do composition vs differential pay of each education-age group can explain the overall gap. In order to separate these components, I conduct a Oaxaca decomposition with agricultural workers as the reference group (Oaxaca (1973)). For notational simplicity, let $F_{st}(education_{ist}, age_{ist}) = \beta_t^s X_{it}^s$, where X_{it}^s is a vector of dummies for each age-education group in sector s. We can then decompose the wage gap in each year as follows:

$$\begin{aligned}\Delta_{s'}(E(log(w_{ist}))) &= \beta_t^a(E(X_{it}^{s'}) - E(X_{it}^a)) + (\beta_t^{s'} - \beta_t^a)E(X_{it}^a) \\ &+ (E(X_{it}^{s'}) - E(X_{it}^a))(\beta_t^{s'} - \beta_t^a)\end{aligned}$$

The first term is entirely due to composition effects due to age and education differences in workers employed by sector s' relative agriculture. In other words, this component reflects the mean wage gap if all education-age groups were equally paid in both agriculture and sector s'. The second term reflects the wage gap due to differential pay of each age and education pair, weighted by the distribution of observable characteristics present in agricultural workers. Unlike the first term, this second component is solely affected by differential returns to age and education, and not by differences in composition. The third term accounts for the interaction between the the composition and return effects.

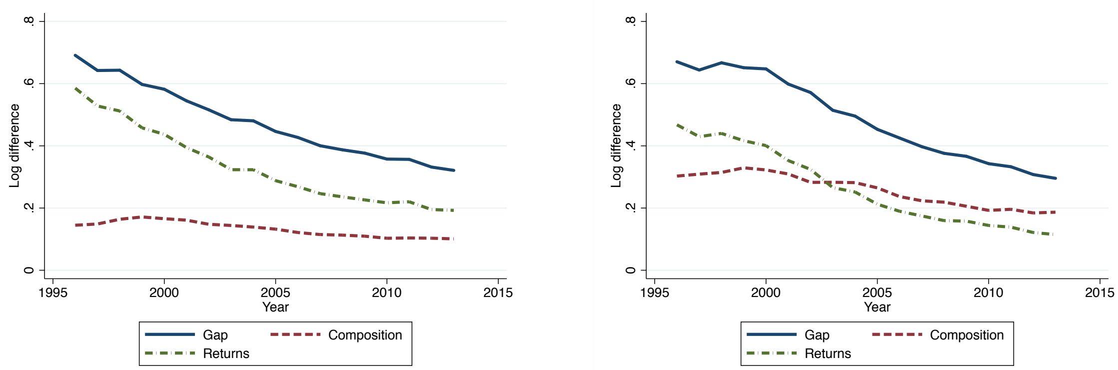
Figure 8 shows the result of this decomposition. Composition effects explain only a small share of the agriculture vs manufacturing gap throughout the sample period, and they explain a larger share, but not all, of the services vs agriculture gap. Differences in the age composition and educational attainment in each sector cannot account for most of the agricultural wage gap in the earlier period, when the gap was largest. Moreover, when looking at the evolution of this decomposition over time, most of the decline in the gap between agriculture and both manufacturing and services is driven by the steeper decline in the gap due to estimated return coefficients.

The limited role of age and education is present in spite of the lack of control for unobservable skill differences between education-age groups. It is likely that this omission overstates the role of compositional differences. For instance, if workers with higher education are paid more not because of their education, but rather because of unobservable skills that are correlated with their education level, this correlation biases upward the share of the wage gap explained by these observable characteristics. Hence, to the extent that more highly paid age-education groups possess more highly valued unobservable skills, the share of the gap explained by observables above is an upper bound on the role of these characteristics. In appendix A, the role of observables after controlling for worker fixed effects is estimated. Since individual workers' changes in age and education have little impact on their wages, controlling for unobservable fixed characteristics erases most of the role of observables in explaining the gap.

Figure 8: Oaxaca decomposition

(a) Agriculture vs Manufacturing

(b) Agriculture vs Services



Note: Gap refers to the difference in mean log wages between two sectors. Returns refer to the term $(\beta_t^{s'} - \beta_t^a)E(X_{it}^a)$ and composition refers to term $\beta_t^a(E(X_{it}^{s'}) - E(X_{it}^a))$ of the Oaxaca decomposition.

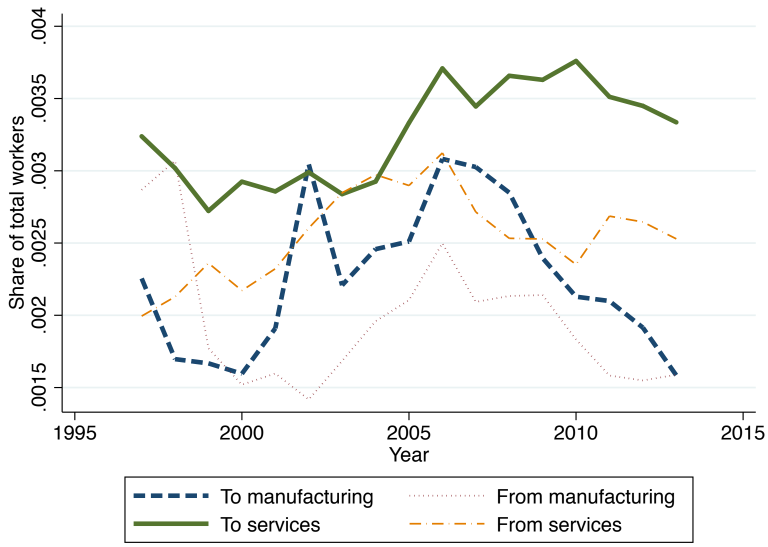
5.2 Unobservable characteristics

The role of differential returns emphasized above does not necessarily imply that workers in agriculture are intrinsically less productive or skilled. There are two types of competing stories that can explain the Oaxaca decomposition above. On the one hand, agricultural workers may have a different composition of unobservable characteristics which makes them less valuable in the market. On the other hand, workers may be similar in the two sectors, but mobility frictions or compensating differentials may induce differential pay for each worker type.

Each of these stories have different implications for the behavior of sector-switches. In the first case, under perfectly competitive labor markets with fully mobile workers, every worker should move to the sector where he or she is paid the most. This process would eliminate any differences in pay among workers with similar —observed and unobserved— characteristics and wage-switchers should not experience large gains. This result is independent of any capital or technological limitations that are particular to each sector. In the second case, compensating differential stories —where workers value sector-specific non-pay characteristics and are therefore willing to receive lower pay in some sectors— or mobility frictions can break this pattern. For instance, one can imagine a situation in which workers are unwilling to pay a mobility cost from moving to industrial areas or one in which workers are unwilling to sacrifice the perks of employment conditions in agriculture. These stories are able to generate wage gaps within each age-education groups that are consistent with the differential returns observed in the previous section and predict that sector-switches should be associated with gains in compensation.

In order to distinguish differential pay from compositional differences in unobservable characteristics, it is necessary to use the panel dimension of the dataset. Using information on sector-switchers, I estimate the magnitude of wage changes from sector transitions controlling for time trends. In order to study these switches, however, enough sector-switchers are needed to estimate these changes precisely. Figure 9 shows the share of workers that switch across sectors throughout the sample period. The small share of sector switchers would usually complicate the study of sector wage jumps using a small-sample panel dataset. However, because of the large number of workers in the sample, this is not a problem. In any given year, there are over ten thousand formal workers who switch into and out of agriculture in the sample.

Figure 9: NUMBER OF SECTOR-SWITCHERS FROM AND INTO AGRICULTURE



Note: Share of total employed workers that switch out of or into agriculture in any given year.

To assess the magnitude of wage changes after controlling for differences in unobserved characteristics, the following worker fixed effect model is estimated

$$log(w_{it}) = \beta_m^t * M_{it}\phi_t + \beta_s^t * S_{it}\phi_t + \phi_t + \phi_i^p + \varepsilon_{it} \quad (1)$$

where M_{it} and S_{it} are indicators for working in the Manufacturing and Services sectors, respectively; ϕ_t and ϕ_i^p are time and individual fixed effects.[9] Individual fixed effects are allowed to vary by six-year periods, but are fixed within each period p.[10] This is done to allow for long-term changes in the distribution of unobservable characteristics. Most importantly, sector indicators are interacted with time; therefore, the coefficients β_s^t and β_m^t reflect average wage changes from switching sectors from agriculture to both manufacturing and services in each year t. I will refer to these coefficients as sector premiums with respect to agriculture, of which there are $2{*}T$ in the model, where T is the number of years in the sample. The model is estimated using all formal workers in Brazil from 1996 to 2013. In the baseline estimation of the model, the sector premiums are identified by workers who switch sectors during this period, and controls are estimated using information from all formal

[9]Since age is collinear with time and individual fixed effects, and education does not change over time for the vast majority of active workers, these controls are not included.

[10]There are three periods in the sample: 1996–2001, 2002–2007, and 2008–2013.

workers in the data.

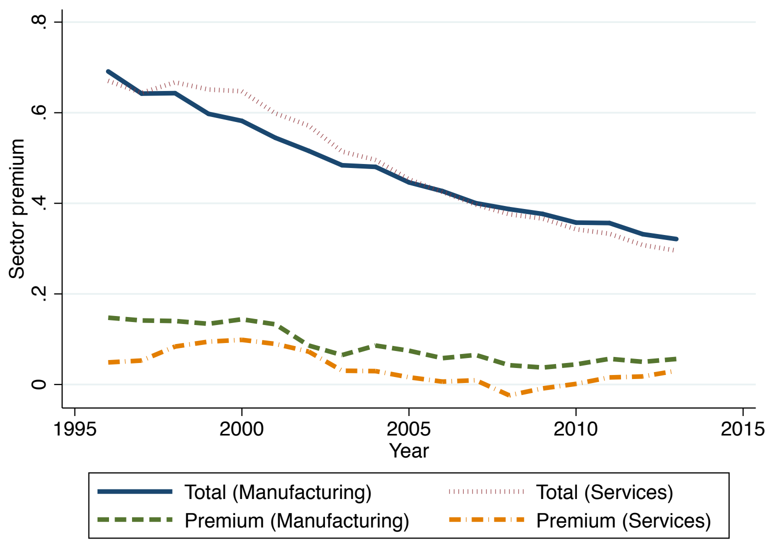
The time series of both services premiums (β_m^t) and manufacturing premiums (β_m^t) are shown in Figure 10. A first takeaway from the figure is that wage differences estimated from switchers are much smaller than the overall wage gap. This is true throughout 1996-2013. For manufacturing, the average sector premium during 1996-2013 is nine log points compared to the overall wage gap of 48 log points relative to agriculture. Similarly, for services, the average jump in wages is four log points compared to the mean total gap of 48 log points. Hence, sector premiums as a percentage of the total gap in a given year averaged 17 percent when comparing agriculture vs manufacturing and seven percent when comparing agriculture to services. Repeating the exercise using earnings instead of hourly wages as a dependent variable provides similar results (Appendix B). The modest magnitude of premium shares suggests that the role of theories producing differential pay of similar workers across sectors is limited.

A key identification assumption of the model is that the error term must be orthogonal to the manufacturing and services dummies. This is violated if workers that switch out of agriculture are precisely the ones who would experience the largest wage jump from switching out of agriculture, which may certainly be the case. In a mobility frictions story, for example, it is precisely the workers who stand to gain the most from transitioning the ones who are willing to overcome this friction and move out of agriculture. Similarly, in a compensation differential story, workers only accept to move out of agriculture if compensated for the loss of non-pay benefits enjoyed in their original sector. These mechanisms, however, would bias our sector premium estimates upwards, so that β_m^t and β_s^t are upper bounds on the potential wage gains to be obtained from switching out of agriculture. To the extent that sector-switchers are the ones who stand to gain the most, this further depresses the role of differential pay stories in explaining the overall wage gap.

Another related concern is that estimates are affected by the inclusion of all workers in the estimation rather than just sector-switchers. Table 2 shows the average sector premium coefficients by period when the model in equation (1) is estimated using only sector-switchers and only transitions out of agriculture. A focus on switchers further lowers the estimates of sector premiums estimated in the baseline for manufacturing, and premiums are similar to the baseline when comparing agriculture to services. Moreover, results do not appear to be driven by asymmetries from sector-switches. This might be a concern if switchers into agriculture are solely driven by improving job offers and these positive job changes counterweight large potential premiums from workers switching out of agriculture. This is not the case, as the model estimated solely on workers who switch out of agriculture yields lower coefficients relative to the baseline for manufacturing, and remain virtually the same for services. In fact, when performing an event-study of workers that exit agriculture (Appendix C), it is not entirely clear whether the wage jump from exiting agriculture is drastically different that the average gain expected from an extra year of experience working in any given sector. There is also no evidence of large improvements in longer term wage growth when comparing pre and post transition trends.

Finally, I have also explored whether there are larger gains from workers switching out the agricultural sector while moving from rural to urban areas at the same time (Appendix D). Moving both out of agriculture and into a city boosts premiums sector premiums by an average of 1–5 log points, still short of the overall gap magnitudes.

Figure 10: Sector gaps relative to agriculture controlling for individual fixed effects



Note: Total refers to the difference in mean log wages between each non-agricultural sector and agricultures. Sector premiums for services (β_s^t) and manufacturing(β_m^t) are defined by equation (1). With the exception of services in 2010, coefficients are all statistically different from zero ($p < .01$).

6 How are compositional differences sustained in equilibrium?

The analysis above suggests that most of the wage gap level —particularly in the late 1990s— is not due to differential pay of equally skilled workers between agriculture and non-

Table 2: Sector premiums relative to agriculture estimated using sector-switchers and all workers

	1996-2001	2002-2007	2008-2013
Services ($\bar{\beta}_s^t$)			
All workers	0.08	0.03	0.01
Sector-switchers	0.10	0.04	0.01
Exiters from agriculture	0.08	0.03	0.01
Manufacturing ($\bar{\beta}_m^t$)			
All workers	0.14	0.07	0.05
Sector-switchers	0.11	0.05	0.04
Exiters from agriculture	0.10	0.05	0.04

Note: Average of sector premiums, $\bar{\beta}_s^t$ and $\bar{\beta}_m^t$, over each six-year period are presented. These are defined by equation (1). All workers category comprise all formal workers between 18 and 65 years old in the RAIS. Sector switchers restrict the sample to workers that have switched into or out of agriculture at least once in each six-year interval. Exiters from agriculture are defined as workers that have switched from agriculture to another sector at least once in each six-year interval.

agriculture. Instead, the wage gap appears to be largely driven by compositional differences in educational attainment and fixed unobservable characteristics between sectors. According to the results presented, a plausible mechanism for generating wage gaps must therefore achieve a very particular goal. It must generate wage gaps driven by large differences in worker characteristics in each sector without giving rise to large differences in pay for similar workers in the two sectors.

Following the work of Roy (1951), recent papers have proposed the sorting of workers with sector-specific skills as a possible explanation of wage and productivity differences between countries, urban vs rural areas, and sectors.[11] This mechanism can generate inter-sector gaps driven by compositional differences in worker characteristics in a manner that is consistent with the empirical observations described. In this section, I test the explanatory power of worker sorting in explaining the wage gap level and its decline. I first assess the existence of sector-specific skills, which is a key assumption of these models. Motivated by this exercise, I then describe and calibrate a sorting model to show how large differences in mean wages between sectors can be generated as an equilibrium outcome of heterogeneous workers with sector-specific skills freely choosing sectors. Finally, I explore potential drivers

[11] See Lagakos and Waugh (2013) and Young (2013).

of the wage gap decline in such model.

6.1 Motivating a Roy Model: The existence of sector-specific skills

A basic premise of Roy models is the existence of occupation or sector-specific skills. In the context of the agricultural wage gap, a worker under this view has agriculture-specific skills and non-agriculture-specific skills, which determine the productivity of the worker when performing sector-specific tasks. By influencing labor productivity, sector-specific skills also determine the potential wage of the worker in each sector and influence his labor allocation decision between sectors. It is not clear, however, whether sector-specific skills exist at all.

Although I cannot test the presence of sector-specific skills directly, I can test an implication of sector-specific skills on wage changes among sector-switchers and the sector premiums described in section 5.2. In particular, I study to what extent are sector premiums driven by workers performing different occupations after switching sectors. In a world where workers have sector-specific skills, wage gains from transitioning out of agriculture into another sector should be more prominent when workers perform a different task in their new sector of employment. If, on the contrary, workers are equally productive regardless of the task performed, then wage changes from transitioning out of agriculture must be driven by other forces that are not necessarily related to an increase in labor productivity.

For example, consider a member of the cleaning staff of an agricultural firm who is considering switching out of agriculture. In a world where sector-specific skills exist, he has the potential to achieve a different level of productivity in the non-agricultural sector. That is, the possibility of performing new tasks (e.g. machinery operation, human-capital intensive tasks) that are fundamentally different from the ones originally performed enable the worker to exhibit sector-specific skills, and therefore improve the productivity of his labor. This change in productivity can in turn induce a wage gain from transitioning sectors. In contrast, if the worker transitions out of agriculture but performs the same set of tasks related to his original cleaning job, we would expect gains to be more limited. Switching sectors without switching occupations limits the realization of sector-specific skills and, therefore, potential wage gains under this view.

I now test whether sector premiums from section 5.2 are significantly reduced once we control for changes in occupation by estimating

$$log(w_{it}) = \gamma_m^t * M_{it}\phi_t + \gamma_s^t * S_{it}\phi_t + \phi^p_{occupation} + \phi_t + \phi_i^p + \varepsilon_{it} \tag{2}$$

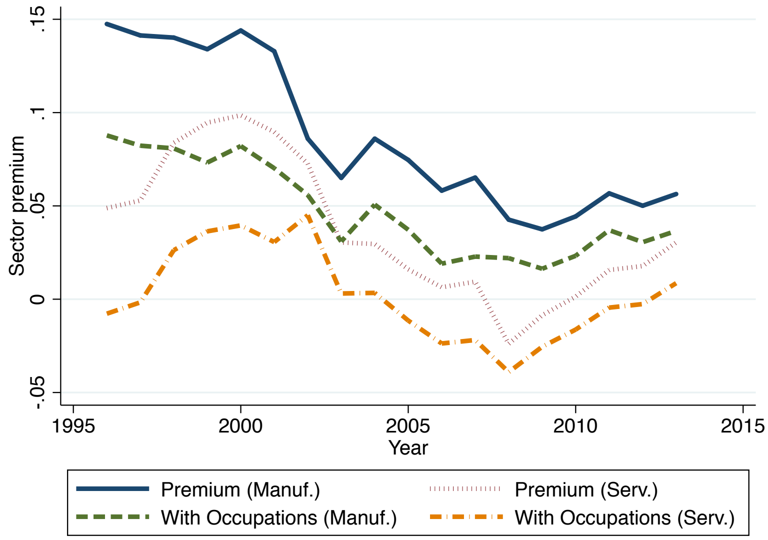
where coefficients γ_m^t and γ_s^t reflect the average differential pay of workers performing the same occupation in both pre and post-transition sectors, $\phi^p_{occupation}$ are occupation fixed effects[12] at the three-digit classification level, and the rest of variables are defined as described in section 5.2. Similarly to the model outlined in the previous section, this model is identified by workers who switch sectors. The main difference of this approach, however, is that the coefficients γ^t_m and γ^t_s are identified using sector-switchers that do not switch occupations after they transition.[13] In the data, several occupations are common to all sectors (e.g. cleaning, security services, drivers/messengers) and the model is therefore identified.

Figure 11 shows the evolution of premiums with (β^t_m, β^t_s) and without (γ^t_m, γ^t_s) occupation controls over time. At the beginning of the period, sector premiums in services disappear after controlling for occupation, and the same is true for the last year of the sample. For manufacturing, accounting for occupational changes reduces the 1996 gap by six log points and the 2013 gap by two log points. The shifting down of both the premium curves after controlling for occupations indicate that a significant portion of these premiums is due to changes in occupation when transitioning sectors. This is consistent with the existence of sector-specific skills which are transformed into wage differences only when performing different tasks in different sectors. To the extent that sector-specific jobs imply the demonstration of sector-specific skills, the downward shift in premiums supports a Roy view of the world where workers have sector-specific abilities.

[12]Occupation fixed effects are allowed to vary by six-year periods but are fixed within the period.

[13]The sector premiums estimated in section 5.2 can be written as $\beta^t_{s'} = \gamma^t_{s'} + (E_{s'}(\phi_{occupation}) - E_{agriculture}(\phi_{occupation}))$ where s is each non-agricultural sector and $E_{s'}(\phi_{occupation})$ is the average of occupation fixed effects in each sector.

Figure 11: Sector premiums after controlling for occupational changes



Note: Sector premiums for services (β_s^t) and manufacturing(β_m^t) are defined by equation (1). Sector premiums controlling for occupations for services (γ_s^t) and manufacturing(γ_m^t) are defined by equation (2). With the exception of services (without occupations) in 2010, premium coefficients are all statistically different from zero ($p < .01$).

6.2 A Roy model of selection with mobility frictions

Motivated by the empirical results shown above, I construct a Roy model to assess the explanatory power of worker sorting in explaining the agricultural wage gap. The following model borrows heavily from Lagakos and Waugh's (2013) framework, deviating from it in three ways. The first is the introduction of a friction that allows for differential pay of similarly skilled workers. The second is allowing changes in the mapping of skill to the marginal productivity of labor, which allows me to explore changes of technology over time. The third is the use of wage information from workers that switch sectors to calibrate the distribution of idiosyncratic and unobservable productivity parameters of workers. In contrast to their framework, the level of the inter sector wage gap is not a calibration target of the model but rather an outcome that I evaluate the model against. In particular, I ask whether a wage gap driven by compositional differences in worker characteristics of the magnitude seen in Brazil can be explained by the worker sorting mechanism. I then assess potential drivers of the wage gap decline in this environment. The model's components are now described in detail.

6.2.1 Preferences and endowments

There is a unit continuum of workers with unit mass, where each worker has identical preferences over agricultural and non-agricultural goods. Following other studies on structural change and the dual economy, workers have a subsistence requirement of $\bar{a}$ agricultural goods. Hence, preferences are non-homothetic and the share of expenditure on agricultural goods increases as income grows.[14] Preferences for each worker i are given by

$$U(c_{a,}^{i}c_{n}^{i}) = log(c_a^i - \bar{a}) + \phi log(c_n^i)$$

where c_a^i and c_n^i refer to consumption of agricultural and non-agricultural goods respectively and ϕ is a weight parameter that determines the relative importance of non-agricultural goods in consumption.

In addition, each worker is endowed with one unit of labor and sector-specific individual skills $\{z_a^i, z_n^i\}$ drawn from a distribution $G(z_a, z_n)$ with support $[\underline{z}, \infty]^2$. A worker can freely choose to work in one of the two sectors but faces a cost k if he decides to work in the non-agriculture sector.[15] This distortionary mobility friction will introduce compensation differences for workers that switch sectors, as observed in the data. In this environment, workers maximize their income y^i, so that $y^i = max\{w_a(z_a^i), w_n(z_n^i) - k\}$ where $w_a(z_a^i)$ and $w_n(z_n^i)$ are the wages offered to worker i in the two sectors. Thus, each worker faces the following budget constraint

$$p_a c_a^i + c_n^i \leq y^i$$

where the non-agricultural good is set to be the numeraire and p_a is the relative price of the agricultural good.

6.2.2 Technology

There is a profit-maximizing representative firm in each sector with production functions given by

$$Y_a = AZ_a, Y_n = AZ_n$$

[14] See Gollin, Parente and Rogerson (2002) and Gollin, Parente and Rogerson (2004) for a more detailed discussion on the role of the subsistence requirement in explaining decreasing agricultural output shares and its role in sector allocations.

[15] Since the focus of the quantitative exercise is analyzing the growth-induced exit of agricultural workers into other sectors, I focus on the mobility cost of going from agriculture to other sectors and not vice-versa.

where A is an economy-wide productivity parameter and Z_s are the total effective units of labor employed by sector s. This second term is equivalent to the sum of individual worker productivities hired by the firm, or $Z_s = \int_{i\epsilon\Gamma^s} z_{is}^{\gamma_s} dG$, where Γ^s is the set of workers hired by sector s and γ_s is a technology parameter governing the mapping of skill to the marginal product of labor in each sector. Similarly, the number of workers employed by sector s is given by $L_s = \int_{i\epsilon\Gamma^s} dG$. Labor productivity of a sector is therefore a function of the integral over the individual production of workers employed in that sector.

6.2.3 Competitive equilibrium

An equilibrium is determined by a relative price of the agricultural good p_a, wage functions $w_a(z_a), w_n(z_n)$, consumption decisions c_a^i, c_n^i and labor allocations such that

1. Firms maximize profits in the two sectors given their technology.
2. Workers' labor allocations maximize their income.
3. Consumption allocations maximize utility subject to the budget constraint.
4. Labor, agricultural, and non-agricultural goods markets clear.

In a competitive market, the first condition requires firms to offer workers a wage equal to the value of their marginal product in their respective sector. These wages vary by both the aggregate productivity factor, relative prices, and the idiosyncratic productivity parameter of each worker.

$$w_a(z_a) = p_a A z_a^{\gamma_a}, w_n(z_n) = A z_n^{\gamma_n}$$

Taking these wages as given, each worker decides to allocate their unit of labor to one of the two sectors. The second condition implies that a worker will chose to work in agriculture as long as the value of their marginal product in agriculture is more than his potential production in non-agriculture minus the mobility cost, or

$$A p_a z_{ia}^{\gamma_a} \geq A z_{in}^{\gamma_n} - k$$

An implication of this is that higher relative agricultural prices, for a given non-agriculture productivity, lowers the minimum agricultural productivity required to stay in agriculture. Moreover, income is given by $y^i = max\{A p_a z_{ia}^{\gamma_a}, A z_{in}^{\gamma_n} - k\}$ and consumption demand for both goods is given by,

$$c_a^i = \frac{y^i}{p_a(1+\phi)} + \bar{a}\frac{\phi}{1+\phi}, c_n^i = \frac{\phi}{(1+\phi)}(y^i - \bar{a}p_a)$$

The above holds as long as $y^i \geq \bar{a}p_a$. Otherwise, $c_a^i = y^i/p_a$ and $c_n^i = 0$. Intuitively, the consumption rule consists of allocating resources on agricultural goods until the minimum subsistence requirement is met and then distributing the remainder among the two goods according to the weight parameter ϕ. Thus, as income grows, a lower proportion of income is allocated to the consumption agricultural goods.

Finally, the market clearing conditions require that good markets clear and that the labor employed in each sector is consistent with workers' labor allocations. This is

$$\begin{aligned}
\int c_n^i dG &= \int_{i \epsilon \Gamma^n} (A z_{in}^{\gamma_n} - k) dG \\
\int c_a^i dG &= \int_{i \epsilon \Gamma^a} A z_{ia}^{\gamma_a} dG \\
\Gamma^n &= \{i : A p_a z_{ia}^{\gamma_a} \leq A z_{in}^{\gamma_n} - k\} \\
\Gamma^a &= \{A p_a z_{ia}^{\gamma_a} > A z_{in}^{\gamma_n} - k\}
\end{aligned}$$

Our main subject of study is the wage gap between agriculture and non-agriculture, as defined in section 5. In the model, this is

$$E(log(w_n)) - E(log(w_a)) = \gamma_n E(log(z_n)) - \gamma_a E(log(z_a)) - log(p_a)$$

The wage gap is therefore the result of three main mechanisms. The first is the direct effect of the relative price, which affects the relative valuation of efficiency units for the output produced in the two sectors. The lower is the relative price of agriculture, the lower is the relative value of agricultural output and hence the greater is the wage gap, holding composition of workers constant. The second mechanism is technology, which changes the average marginal products of workers in the two sectors. The third mechanism is selection, which affects the skill distribution of the sets of workers working in the two sectors (Γ^a, Γ^n). The lower is the relative price of the agricultural good, the more people exit agriculture to work in the other sector. This process changes the composition of workers in each sector, which can increase or decrease the gap in mean worker productivity between sectors. The wage gap level is therefore larger or smaller depending on the equilibrium effect of these mechanisms. Moreover, as a country grows richer and the relative price of agriculture declines[16], the net effect of both price and composition effects on the wage gap over time is undetermined. Whether wage gaps decline or rise in this environment depends on the parameters of the economy.

6.2.4 Calibration strategy

I now proceed to calibrate the model to conduct two types of exercises. The first is testing whether worker sorting can generate wage gaps that are of the same order of magnitude relative to the gaps in the data. The second is exploring potential drivers of the gap decline in a worker sorting environment.

To do this, I first calibrate the economy to Brazil in the earliest period of 1996-1997, when the gap was the largest. Later, growth in parameter A is introduced to match real output growth rates observed in 1996-2013. Preference, production, friction and productivity distribution parameters are jointly estimated to match different moments of the data. Although all of these parameters interact in the model, each of them has stronger implications for particular moments. Below, I describe the relationship of each parameter to each moment and how these are calibrated.

Preference parameters Consistent with the literature using dual economy models with minimum subsistence requirements, preference parameters ϕ and $\bar{a}$ are calibrated to match two moments of the data that relate to labor and output shares. The first is the share of workers in agriculture of 25 percent observed in 1996. The second is a long-run agriculture output share of 0.5 percent, which is the standard parameter used by Lagakos and Waugh (2013), Restuccia, Yang and Zhu (2008) and other studies on structural change. Once calibrated, the minimum subsistence requirement is ten percent of the average wage.

Production and friction parameters The technology parameter A is set to one for the initial calibration in 1996. Later, when studying the effect of growth on the wage gap, changes in A are calibrated to match total yearly real output growth in Brazil during 1996-2013. The sector-neutral nature of the productivity parameter in this economy implies that the wage gap is solely dependent on endogenous price and selection effects in the model.[17]

[16] As a country grows richer, non-homothetic preferences imply that a lower share of income is allocated to agricultural goods consumption and a lower relative price of agricultural goods. See Lagakos and Waugh (2013) for a detailed discussion.

[17] Alternatively, one could introduce distinct growth rates by sector. The impact of differential growth on the wage gap, however, is similar to the one described in this paper. Sector-specific growth rates change the relative price of agriculture which induces the exit of workers out of agriculture and changes worker

The technology parameters γ_a and γ_n are set to one in 1996 so that there is a linear mapping from skill to output in the baseline, as in Lagakos and Waugh (2013). This assumption is relaxed when studying the role of skill-biased technological change discussed below. Finally, the mobility friction k is calibrated to match the average wage gain from switching out of agriculture in the model to the weighted average of the non-agriculture sector premiums for 1996. The relative small transition wage gains relative to the overall gap described in section 5.2 imply that the magnitude of k is relatively small at 14 percent of the minimum subsistence requirement and 1.4 percent of the average wage in the economy. This confirms the conclusion in the previous section: micro-data is not consistent with mobility frictions generating large pay differentials in sector-switching workers.

The joint distribution of sector-specific worker skills I calibrate the distribution parameters using information from wage dispersion within each sector, as well as information from workers that have worked in two sectors. First, I restrict $\underline{z}$ to be equal to k, so that every worker in the economy has a labor productivity endowment that is sufficient to afford a transition out of agriculture. That is, workers close to $\underline{z}$ will freely choose to stay in agriculture because of preferences, and not because of the lack of sufficient income potential to pay the cost of moving. In this way, the wage gap is entirely due to endogenous selection and prices, and not to the distribution's support parameter.

Since I do not observe all workers at all sectors, it is impossible to calibrate the distribution G without imposing some structure. Non-parametric estimation is therefore not an option. Instead, following Lagakos and Waugh (2013), I allow workers to have dependent draws from sector-specific Fréchet distributions $X(z_a)$ and $Y(z_n)$ and restrict the joint distribution $G(z_a, z_n)$ to be a Frank copula resulting from the two primary distributions,

$$
\begin{aligned}
G(z_a, z_n) &= C[X(z_a), Y(z_n)] \\
C[u, v] &= -1/\rho * log(1 + \frac{(e^{-\rho u} - 1)(e^{-\rho v} - 1)}{e^{-\rho} - 1}) \\
X(z_a) &= e^{-z_a^{-\theta_a}} \\
Y(z_n) &= e^{-z_n^{-\theta_n}}
\end{aligned}
$$

The sector-specific distributions have dispersion parameters θ_a and θ_n, which control the

productivity composition in each sector. Regardless of whether growth is sector-neutral or not, the wage gap level is solely determined by differences in the composition of workers between sectors.

within sector variance of the productivity distribution in agriculture and non-agriculture. These are calibrated to match the variances of log wages in agriculture and non-agriculture of 0.52 and 0.77 respectively. The calibrated parameters that match these log variances are $\theta_a = .72$ and $\theta_n = .53$.

Besides the transparent mapping that exists between these parameters and wage dispersion, there are two other reasons why a Fréchet shape is a sensible choice to model sector-specific distributions. First, the Fréchet distribution is a special case of the extreme value distribution; therefore, the marginal Fréchet distribution of a particular sector can be interpreted as the distribution of the maximum draw from a set of productivity distributions within that sector. For example, this may represent the maximum productivity draw out of a series of jobs that are available within the manufacturing or agricultural sectors.[18] Second, the shape of of the distribution, with greater mass at lower productivity parameters and fat tails, resembles the within sector distribution of both raw wages and unobservable worker characteristics observed in the data.

To form a joint distribution out of the two sector-specific marginal distributions, a Frank copula is used. The advantage of using this copula is that it allows the degree of dependence in the two distributions to be controlled by a single parameter ρ. Along with θ_a and θ_n, this parameter is calibrated to match the fraction of workers that switch sectors during 1996-1997 (1 percent of all workers) when the economy grew 2.2 percent in real terms. Intuitively, for given dispersion parameters (θ_a, θ_n), ρ controls the amount of workers close to the labor allocation indifference condition ($Ap_a z_a^i = Az_n^i - k$). A growth-induced change in prices pushes a larger or smaller share of workers out of agriculture depending on the mass of workers that are close to indifferent in the base year. The resulting parameter from this calibration is $\rho = 8$, which implies a linear correlation of 34 percent between the sector-specific productivity parameters z_a and z_n. Importantly, no difference in mean productivity between sectors is assumed in the calibration of the joint distribution. The agricultural wage gap is therefore not a calibration target but an outcome of the model.

6.3 The explanatory power of the sorting mechanism

The model generates both cross-sectional and inter-temporal predictions. In the 1996 cross-section, the predicted wage gap by the calibrated model is 73 log points. This magnitude is five log points higher than the wage gap observed. Importantly, this gap is not the result of the mobility friction k, which equals 0.1 percent of the overall gap and 1.4 percent of the

[18]By the extreme value theorem, the maximum of independent draws from any distribution converges to an extreme value distribution. The Fréchet is an example of these distributions.

average wage. Re-estimating the model with $k = 0$ yields a wage gap of 61 log points (90 percent of the total wage gap observed). Overall, the results indicate that large wage gaps of the magnitudes observed can be generated by compositional differences in skills arising from a sorting equilibrium.

6.3.1 What drove the decline?

Once the economy is calibrated to 1996, I examine potential drivers of the wage gap decline. In particular, I explore whether growth, technological change, or a compression in worker skills can drive the declining pattern in the data. For each potential driver, the model produces a time series of predicted wage gaps (Figure 12) as well as predicted agricultural labor shares (Figure 13).

Economic growth Growth is here introduced as an increase in the sector-neutral productivity parameter A. Changes in this parameter are calibrated to match real GDP growth throughout the period.

Economic growth has two effects in the model. On the one hand, higher income decreases the demand for agricultural goods consumed in the economy relative to non-agricultural goods. This is a direct consequence of the subsistence requirement present in preferences. The reduction in relative demand induces a lowering of the relative price of the agricultural good, which decreases the market value of the marginal product of agricultural workers. This in turn depresses relative wages in agriculture and widens the gap. On the other hand, the price changes cause the exit of workers out of agriculture. The transitioning workers can lower the average skill of workers in the non-agriculture sector, resulting in a decline in the agricultural wage gap. As mentioned before, the net effect of these two forces depends on the economy's parameters.

The evolution of the share of agricultural workers and the agricultural wage gap is shown in dashed blue lines in Figures 12–13. The model fails to generate a fall in the wage gap, which increases by five log points from 1996 to 2013. The increase in aggregate productivity does decrease the share of workers employed in agriculture from 25 to 14 percent. This is in line with the decline seen in the data. However, the recomposition of worker skills in each sector due to the exodus from agriculture fails to produce a wage gap decline. Similar results are obtained if one introduces growth exclusively in the agricultural sector.[19]

[19] Higher productivity growth in the agricultural sector reduces demand for agricultural goods and agricultural workers. This depresses wages in agriculture, resulting in an agricultural wage gap increase.

Skill-biased technological change Technological change has been proposed as a driving force behind wage dynamics and the exit of workers out of agriculture.[20] In Brazil, Bustos, Caprettini and Ponticelli (2016) argue that there have been both labor-saving and labor-augmenting technological innovations in agriculture that have shifted the marginal products of workers in the agricultural sector. As mentioned above, changes in the aggregate productivity of agricultural labor cannot account for the decrease in the wage gap in the model. Nonetheless, skill-biased technological change can produce a decline.

To explore this story, we depart from the original assumption of a linear mapping from skills to the marginal product of labor ($\gamma_a = \gamma_n = 1$ in the baseline). Instead, we let the marginal product of labor to be dependent on the level of skill by letting γ_a and γ_n to vary over time. Since I do not have direct information about the changing technology of firms in Brazil, these parameters are calibrated to match the within-sector variance of log wages. In the calibration, γ_a decreases from one to 0.85 and γ_n decreases from one to 0.90 from 1996 to 2013. The lower values indicate a flatter mapping of skill differences to marginal products and, therefore, to wages.

The dashed-dotted green lines in Figures 12–13 show the predicted wage gaps and agricultural labor shares, respectively. Skill-biased technological change induces a decline in the wage gap from 72 to 60 and relatively stable shares of workers in agriculture around 25–26 percent. When we allow for both technological change and growth in the productivity parameter A (dashed green lines), the wage gap also declines (from 72 to 60 log points) and the labor force shares in agriculture decline to 14 percent.

The results suggest that skill-based technological change can explain part of a decline in the agricultural wage gap. However, the magnitudes of the predicted decline (12 log points) account for only 32 percent of the 37 log point decline observed in the data. Moreover, the changes in the mapping of skills to marginal products is at odds with commonly proposed stories of skill-biased technology. In particular, the model suggests that technology has flattened the mapping from skills to output by increasing the marginal product of low-skilled workers relative to high-skilled ones. A labor-saving mechanization story would have the opposite effect, and it is unclear whether improvements in the use of intermediate goods (e.g. use of improved seeds or pesticides) have this type skill-biased effects. For skill-biased technological change to have contributed to the wage gap decline, the drivers of such change must have closed the gaps in marginal products between low and high skilled workers instead

[20] Ngai and Pissarides (2007) and Baumol (1967) emphasize the connection between productivity growth, prices and the exit of workers out of agriculture. The connection between skill-biased technological change and wage differences has been explored in the United States. See Autor, Katz and Kearney (2008) for a review.

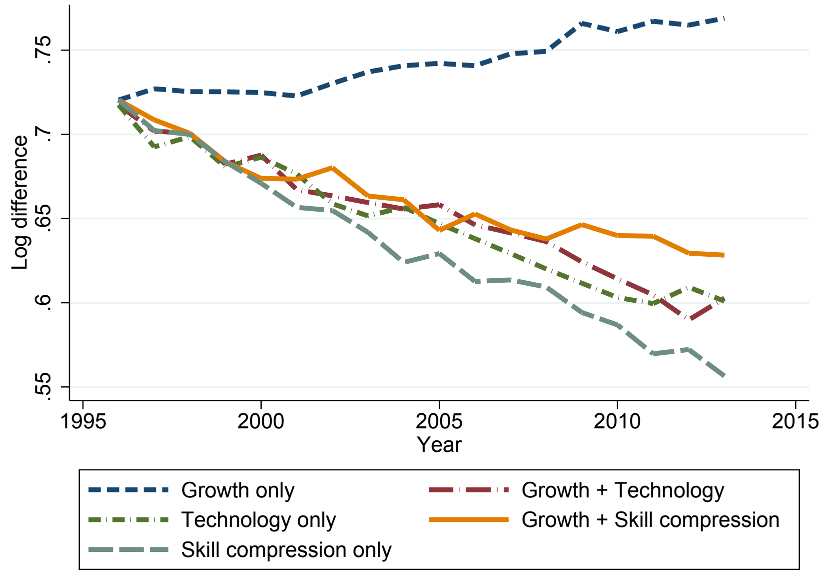
of widening them.

Compression of the skill distribution Section 5 emphasized the role of worker observable and unobservable skills in explaining the wage gap level and its decline. I now explore whether a compression in worker heterogeneity can lead to a compression of the average wage gap. To do this, I start with the calibrated 1996 economy as a baseline, and let the skill dispersion parameters θ_a and θ_m to vary over time. In particular, these parameters vary to match the changes in within-sector wage variance. The within-agriculture variance decline in log earnings from 0.44 in 1996 to 0.23 in 2013 results in an increase in θ_a from 0.73 to 1.16. Similarly, the variance in non-agriculture declines from 0.77 to 0.50 increases θ_n from 0.52 to 0.73.

The dashed gray line in Figure 12 shows the gaps generated by the compression in the distribution of worker skills. There is a reduction of the gap of 17 log points, representing 46 percent of the total gap decline in the data. A shortcoming of this exercise, however, is that without any growth in the productivity parameter A, we do not see a decrease in the share of workers in agriculture (Figure 13). An alternative and arguably more realistic exercise is to allow for both the productivity parameter and the distribution of skills to vary over time. This exercise is shown in solid orange lines in Figures 12–13. When allowing for both growth and a compression of skill characteristics, the gap decreases by nine log points, which is 24 percent of the observed decline. The share of workers in agriculture drops to 21 percent over the period.

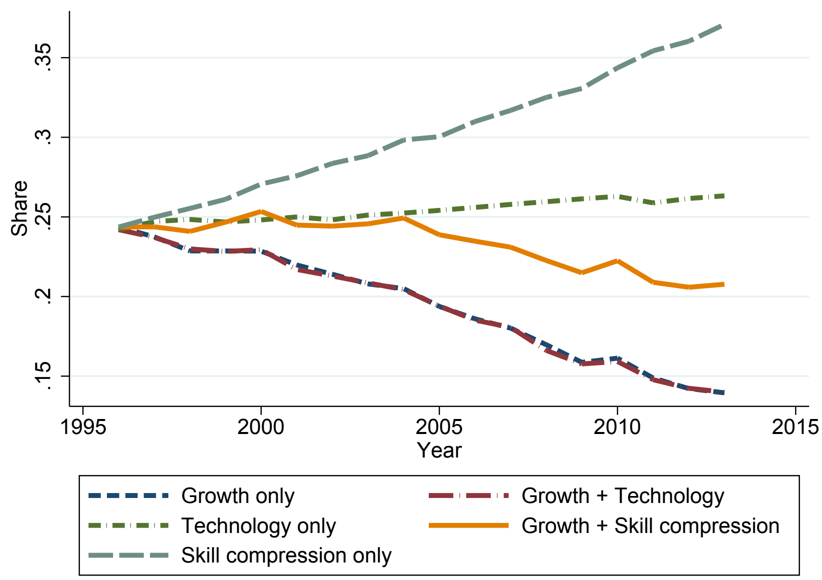
There are two main takeaways from this analysis. First, growth accompanied by a compression in the skill distribution can lead to qualitative drop in both the share of workers in agriculture and the average wage gap. Second, neither of these forces can quantitatively account for the total magnitude of the declines in both the wage gap and the agricultural labor share.

Figure 12: Agricultural wage gaps



Note: Wage gap defined as the difference in the mean of log wages in agriculture and non-agriculture.

Figure 13: Share of workers in agriculture



7 Conclusion

The large wage gaps between agriculture and other sectors are likely driven by compositional differences in worker characteristics. In accordance with cross-country patterns on inter-sector productivity and wage gaps, these differences declined gradually in Brazil since the late 1990s as the country became richer. This paper argues that a Roy model of worker sorting based on sector-specific skills is consistent with the magnitude of the wage gaps observed. Moreover, a reduction of disparities in worker productivity in a worker sorting environment can lead to a decline over time as workers transition out of agriculture.

Nonetheless, a significant share of the wage gap decline remains unexplained, and it is unclear what drove the compression of within sector heterogeneity in worker productivity. On the one hand, the skill-based technological change required to reduce gaps is not easily reconcilable with common narratives of agricultural mechanization and intermediate input usage. On the other hand, little of the compression in gaps can be explained by education composition, though compression in educational quality remains a plausible driver.

Exploration of additional mechanisms is therefore needed, and the evidence presented in this paper can guide the design of future wage gap models. Specifically, the results show that pay differences for workers with similar skills are relatively small when compared with the total wage gap. This finding discourages models that generate large pay differences for similar work in different sectors. For instance, mobility frictions or compensation differentials that induce large gaps in wages per efficiency units between sectors would predict large wage gains from switching sectors. These are at odds with the data. Complementary mechanisms to worker sorting that attempt to rationalize the wage gap level and its decline must, at the very least, produce large differences in average pay between sectors without producing relatively large wage gains for workers that exit agriculture.

References

Adamopoulos, Tasso and Diego Restuccia. 2014. "The Size Distribution of Farms and International Productivity Differences." *American Economic Review* 104(6):1667–1697.

Autor, David H., Lawrence Katz and Melissa Kearney. 2008. "Trends in US wage inequality: Revising the revisionists." *The Review of Economics and Statistics* 90(2):300–323.

Baumol, William. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of the Urban Crisis." *American Economic Review* 57:415–426.

Beegle, Kathleen, Joachim De Weerdt and Stefan Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* 93:1010–1033.

Bryan, Gharad, Shyamal Chowdhury and Ahmed Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica* 82(1671-1748).

Bustos, Paula, Bruno Caprettini and Jacopo Ponticelli. 2016. "Agricultural Productivity and Structural Transformation: Evidence from Brazil." *American Economic Review* 100(6):1320–1365.

Caselli, Francesco. 2005. "Accounting for Cross-country Income Differences." *Handbook of economic growth* 1:679–741.

Chiquiar, Daniel and Gordon Hanson. 2005. "International Migration, Self-Selection, and The Distribution of Wage: Evidence from Mexico and the United States." *Journal of Political Economy* 113(239-281).

Donovan, Kevin. 2016. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences.".

Gibbons, Robert and Lawrence Katz. 1992. "Does Unmeasured Ability Explain Inter-industry Wage Differentials?" *The Review of Economic Studies* 59(3):515–535.

Gollin, Douglas, David Lagakos and Michael Waugh. 2014. "Agricultural Productivity Gap in Developing Countries." *Quarterly Journal of Economics* 129(2):939–993.

Gollin, Douglas, Stephen L Parente and Richard Rogerson. 2004. "Farm Work, Home Work and International Productivity Differences." *Review of Economic Dynamics* 7(4):827–850.

Gollin, Douglas, Stephen Parente and Richard Rogerson. 2002. "The Role of Agriculture in Development." *American Economic Review* pp. 160–164.

Hendricks, L. and Todd Schoellman. 2017. "Human Capital and Development Accounting: New Evidence from Wage Gains at Migration." Working paper.

Herrendorf, Berthold and Arilton Teixeira. 2011. "Barriers to Entry and Development." *International Economic Review* 52(2):573–602.

Herrendorf, Berthold and Todd Schoellman. 2015. “Why is Measured Productivity So Low in Agriculture?” *Review of Economic Dynamics* 18(4):1003–1022.

Hicks, Joan Hamory, Marieke Kleemans, Nicholas Li and Edward Miguel. 2017. “Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata.”.

Klenow, Pete and Chang-Tai Hsieh. 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics* 124:1403–1448.

Krueger, Alan B and Lawrence H Summers. 1988. “Efficiency Wages and The Inter-Industry Wage Structure.” *Econometrica: Journal of the Econometric Society* pp. 259–293.

Kuznets, Simon. 1971. *Economic Growth of Nations: Total Output and Production Structure*. Cambridge: Harvard University Press.

Lagakos, David and Michael E Waugh. 2013. “Selection, Agriculture and Cross-country Productivity differences.” *American Economic Review* 103:948–980.

Lewis, W Arthur. 1955. *Theory of Economic Growth*. Routledge.

Murphy, Kevin M and Robert H Topel. 1987. “Unemployment Risk and Earnings.” *Unemployment and the Structure of Labor Markets* .

Murphy, Kevin M and Robert H Topel. 1990. “Efficiency Wages Reconsidered: Theory and Evidence.” *Advances in the Theory and Measurement of Unemployment* pp. 204–240.

Ngai, L. Rachel and Christopher A. Pissarides. 2007. “Structural Change in a Multisector Model of Growth.” *American Economic Review* 97:429–443.

Oaxaca, Ronald. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” *International Economic Review* 14(3):693–709.

Restuccia, Diego, Dennis Tao Yang and Xiaodong Zhu. 2008. “Agriculture and Aggregate Productivity: A Quantitative Cross-country Analysis.” *Journal of Monetary Economics* 55(2):234–250.

Restuccia, Diego and Richard Rogerson. 2008*a*. “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments.” *Review of Economic Dynamics* 11(4):707–720.

Restuccia, Diego and Richard Rogerson. 2008*b*. “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments.” *Review of Economic Dynamics* 11(4):707–720.

Rostow, Walt Whitman. 1960. *The Stages of Economic Growth: A Non-communist Manifesto*. Cambridge University Press.

Roy, Andrew Donald. 1951. "Some Thoughts on The Distribution of Earnings." *Oxford economic papers* 3(2):135–146.

Vollrath, Dietrich. 2014. "The Efficiency of Human Capital Allocations in Developing Countries." *Journal of Development Economics* 108:106–118.

Yang, Dean. 2006. "Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants' Responses to Exchange Rate Shocks." *Review of Economics and Statistics* 88(4):715–735.

Young, Alwyn. 2013. "Inequality, the Urban-Rural Gap and Migration." *The Quarterly Journal of Economics* .

A The gap explained by observable differences after controlling for unobservable characteristics

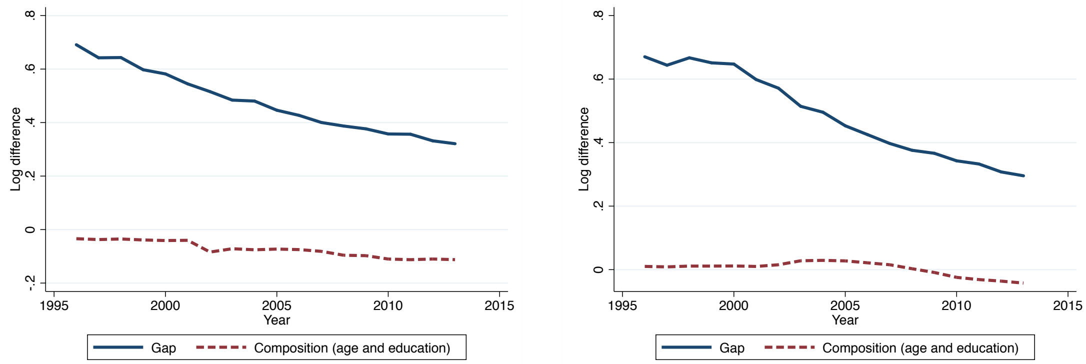
Differences in pay across sectors for each age and education group can potentially be explained by differences in the composition of unobservables. It might be the case that, for a a given age and education level, workers in one sector are fundamentally more productive than others. Workers certainly differ in dimensions other than age and education, and the literature has emphasized the importance of unobservable skill heterogeneity in explaining wage differences. In order to analyze the relative importance of observed and unobserved worker characteristics, I estimate the following model using the panel dimension of the dataset for each sector.

$$log(w_{ist}) = \phi^{p}_{si} + F_{st}(education_{ist}, age_{ist}) + \varepsilon_{ist} \tag{3}$$

where ϕ^{p}_{si} is the fixed effect of a worker i in sector s, which is allowed to vary by each six year period p. This specification controls for worker unobservables that are fixed over time when estimating differential compensation of age and education levels. As before, $F_{st}(education_{ist}, age_{ist})$ is a sum of education, age and year interacted dummies. Unlike the baseline specification in section 5, the identification of age and education returns comes from the workers that switch age and education categories throughout the period of study. Figure 14 shows the role of age and education differences for both manufacturing and services under this specification. When individual fixed effects are included, differential composition of education and age across sectors explains nearly none of the gap.

Figure 14: Oaxaca decomposition controlling for unobservables

(a) Agriculture vs manufacturing (b) Agriculture vs services

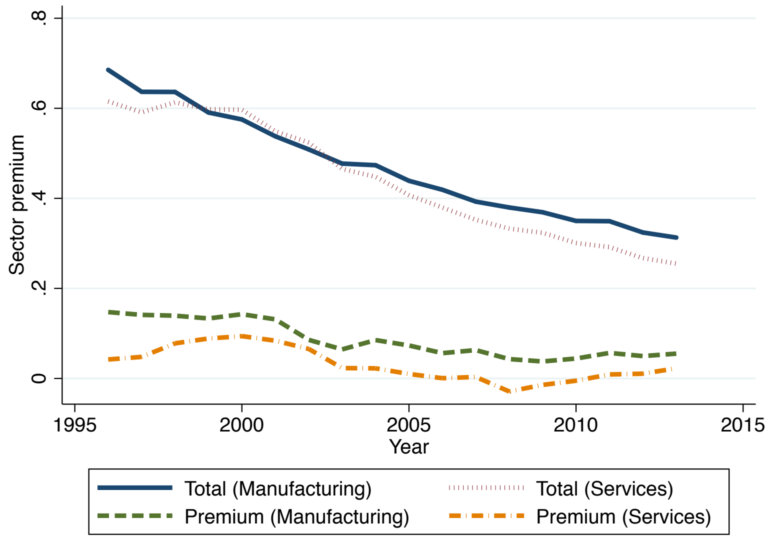


Note: Gap refers to the difference in mean log wages between sectors. Age and education refer to the difference of the mean predicted values, $E(F_{st}(education_{ist}, age_{ist})|s = a) - E(F_{st}(education_{ist}, age_{ist})|s = a)$.

B Sector premiums using earnings

The analogous results from Figure 10 using monthly earnings instead of hourly wages as a dependent variable are shown in Figure 15. Similarly to the baseline, earnings premiums are smaller than the overall earnings gap. For manufacturing, the average sector premium during 1996-2013 is 11 log points compared to the overall earnings gap of 48 log points relative to agriculture. Similarly, for services, the average jump in earnings is eight log points compared to the mean total gap of 46 log points. Sector premiums as a percentage of the total gap in a given year averaged 21 percent when comparing agriculture vs manufacturing and 16 percent when comparing agriculture to services.

Figure 15: Sector gaps in earnings relative to agriculture controlling for individual fixed effects



Note: Total refers to the difference in mean log wages between each non-agricultural sector and agricultures. Sector premiums for services (β_s^t) and manufacturing(β_m^t) are defined by equation (1). With the exception of 2008, coefficients are all statistically different from zero ($p < .01$).

C Event-study of transitions into and out of agriculture

This section adopts an event-study framework focusing on workers that switch out of agriculture. The following equation is estimated for transitions out of agriculture into both services and manufacturing.

$$log(w_{it}) = \sum_{j=-2}^{5} \gamma_j + \phi_t + \phi_i + \varepsilon_{it}$$

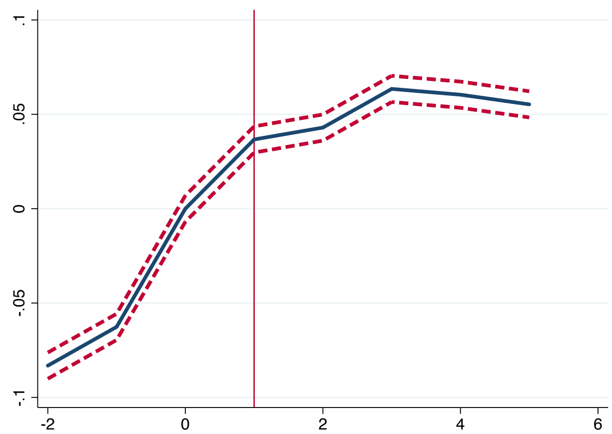
As before, ϕ_t are year effects and ϕ_i are worker fixed effects. In order to consider a longer timespan of transitions, fixed effects are not allowed to vary by period. Coefficients γ_j are dummy indicators for pre and post transition years. These coefficients are only equal to one if a worker is observed three-years before transitioning and five years after transitioning out of agriculture. This is done in order to avoid selection effects in the estimation of transition coefficients. All workers are included in this exercise in order to better estimate year effects.

Figure 16 shows the results of the transition coefficients, γ_j, with confidence intervals.

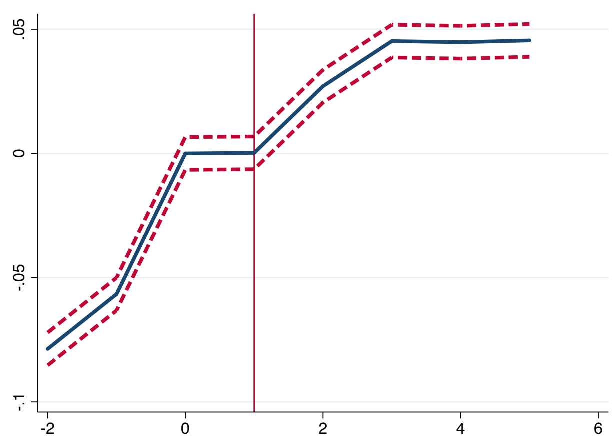
Transitions into services and manufacturing are analyzed separately with similar results. As before, the wage increases five years after transitions are much smaller than the magnitude of the aggregate wage gap. Furthermore, there is no evidence of improved wage growth profiles after transitioning out of agriculture. If anything, there is a flattening of the wage growth profile after transitioning into the non-agriculture sector. Figure 17 shows similar results when age squared is added as an additional control.

Figure 16: Transitions out of agriculture

(a) From agriculture to manufacturing



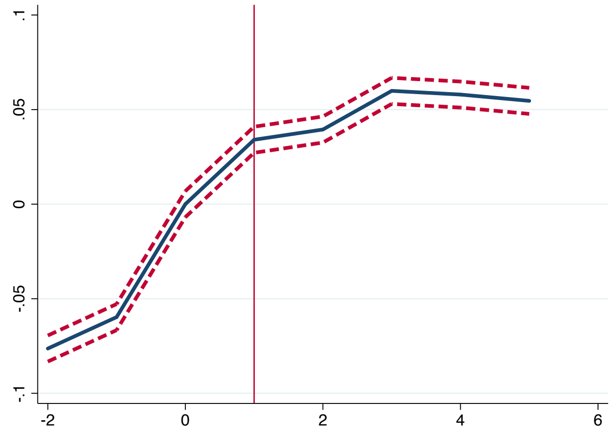
(b) From agriculture to services



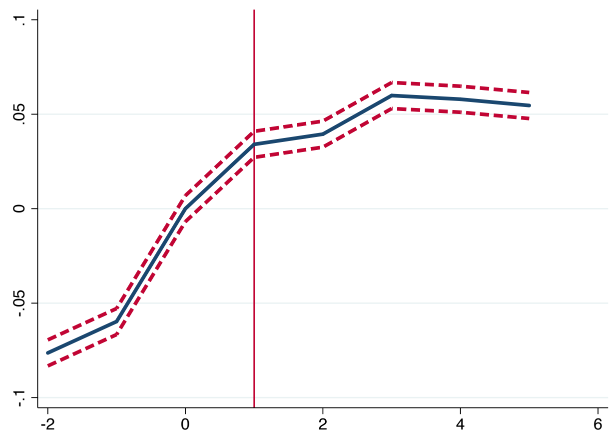
Note: Year 0 refers to the last year worked in the pre-transition sector and Year 1 refers to the first year in the post-transition sector. Solid line shows coefficients γ_j subtracted by γ_0 so that coefficients reflect changes relative to the pre-transition wage level. Dashed lines are 95 percent confidence intervals from transitions after controlling for year effects.

Figure 17: Transitions out of agriculture controling for age squared

(a) From agriculture to manufacturing



(b) From agriculture to services



Note: Transition coefficients controlling for year effects and age^2. Year 0 refers to the last year worked in the pre-transition sector and Year 1 refers to the first year in the post-transition sector. Solid line shows coefficients γ_j substracted by γ_0 so that coefficients reflect changes relative to the pre-transition wage level. Dashed lines depict 95 percent confidence intervals from transitions after controlling for year effects.

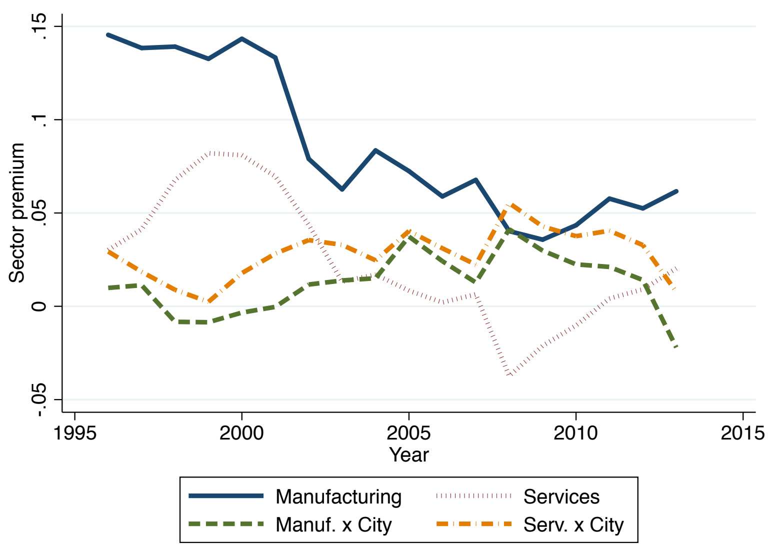
D Transitions into cities

This section explores whether sector premiums are much larger for workers that not only transition sectors but also transition into urban areas. To explore this, the following specification is estimated:

$$log(w_{it}) = \beta_m^t * M_{it}\phi_t + \beta_s^t * S_{it}\phi_t + \delta_m^t * city_{it} * M_{it}\phi_t + \delta_s^t * city_{it} * S_{it}\phi_t + \lambda^t * city_{it} * \phi_t + \phi_t + \phi_i^p + \varepsilon_{it}$$

where M_{it} and S_{it} are indicators for working in the Manufacturing and Services sectors, respectively; ϕ_t and ϕ_i^p are time and individual fixed effects (which vary by period); and $city_{it}$ is an indicator for working in a city. A city is defined as a municipality with more than one hundred thousand formal workers in the dataset. There are 37 out of 5,570 municipalities that classify as cities under this definition. This specification allows for the decomposition of sector premiums into the ones estimated from sector transitions alone, β_s^t and β_m^t, and the sector premiums associated to transitions between sectors into cities, δ_s^t and δ_m^t. The evolution of these coefficients over time is shown in Figure 18. The coefficients on the city and sector interactions average 1–5 log points while sector transitions that occur without changes in city/non-city status average premiums of 2–9 log points throughout the period. Thus, he results show only moderate additional compensation gains from switching into urban areas on top of switching sectors.

Figure 18: Sector premiums with transitions into cities



Note: Sector premiums for services (β_s^t), manufacturing(β_m^t), services x city (δ_s^t) and manufacturing x services (δ_m^t). A municipality is defined to be a city if it has over one hundred thousand formal workers.