



WP/17/289

IMF Working Paper

Structural Transformation and the Agricultural Wage Gap

by Jorge A. Alvarez

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I N T E R N A T I O N A L M O N E T A R Y F U N D

IMF Working Paper

Research Department

Structural Transformation and the Agricultural Wage Gap

Prepared by Jorge Alvarez*

Authorized for distribution by Chris Papageorgiou

December 2017

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

A key feature of developing economies is that wages in agriculture are significantly below those of other sectors. Using Brazilian household surveys and administrative panel data, I use information on workers who switch sectors to decompose the drivers of this gap. I find that most of the gap is explained by differences in worker composition. The evidence speaks against the existence of large short-term gains from reallocating workers out of agriculture and favors recently proposed Roy models of inter-sector sorting. A calibrated sorting model of structural transformation can account for the wage gap level observed and its decline as the economy transitioned out of agriculture.

JEL Classification Numbers: J24, J31, J43, J62, Q10, R10.

Keywords: Wage Gaps, Structural Transformation, Agriculture, Human Capital, Sorting, Productivity Gaps, Brazil

Author's E-Mail Address: jalvarez@imf.org

* I am grateful for the insightful and generous advice of Richard Rogerson throughout 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, Rafael Dix-Carneiro, 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.

Contents

1 Introduction	4
2 Literature Review	6
3 Data description	9
4 The magnitude and evolution of the agricultural gap in Brazil	12
5 Sources of the agricultural gap	15
5.1 Worker observables and human capital	16
5.2 Unobservable characteristics	19
6 How are compositional differences sustained in equilibrium?	25
6.1 Motivating a Roy Model: the existence of sector-specific skills	25
6.2 A Roy model of selection with mobility frictions	28
6.2.1 Preferences and endowments	29
6.2.2 Technology	30
6.2.3 Competitive equilibrium	30
6.2.4 Calibration strategy	32
6.3 The explanatory power of the sorting mechanism	35
6.3.1 What drove the decline?.....	35
7 Conclusion	39
References	41
Appendix	
A The decline in wage gaps throughout the distribution	45
B The evolution of output gaps in Brazil.....	46
C The evolution of agricultural labor shares.....	48
D Sector-switchers over time.....	48
E Sector premiums using earnings.....	49
F Event-study of transitions into and out of agriculture	50
List of Figures	
1 Wage gap in Brazil.....	15
2 Residual gap after controlling for observables	17
3 Oaxaca decomposition	19
4 Sector gaps relative to agriculture controlling for individual fixed effects	23
5 Sector premiums with transitions into cities	24
6 Sector premiums after controlling for occupational changes	28
7 Agricultural wage gaps	38
8 Share of workers in agriculture.....	39

9	Gaps in Brazil by percentile.....	46
10	Value added per worker gap in Brazil	47
11	Workers by sector	48
12	Number of sector-switchers from and into agriculture	49
13	Sector gaps in earnings relative to agriculture controlling for individual fixed effects...	50
14	Transitions out of agriculture.....	51
15	Transitions out of agriculture controlling for age squared.....	52

List of Tables

1	RAIS Summary Statistics	13
2	Sector premiums relative to agriculture estimated using sector-switchers and all workers	24

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 earn different wages, it is hard to distinguish whether the two sectors have differential pay for similar workers or whether the two workers have different unobserved characteristics. To make this distinction, the following strategy is adopted. First, using a collection of household surveys on formal and informal workers, I show that differences in age, education, gender, race and regional controls account for part, but not most, of the observed wage gap in Brazil. Second, using administrative panel data on the universe of formal workers, I show a quantitatively and qualitatively similar pattern. Third, I use the panel dimension of the administrative data to assess the role of unobserved characteristics. 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 use of structural models to match country-level moments or on the analysis of heterogeneous cross-sectional surveys

¹In 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.

from a sample of countries. With panel data, 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.

This study finds that formal workers who transition out of agriculture experience limited compensation gains when compared to the large overall gap in mean wages between agriculture and other sectors. I conclude that the agricultural wage gap in the formal sector is not driven by differential pay of similar workers, once fixed unobservable characteristics are controlled for. Instead, most 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 declined significantly from 1996 to 2013 as the economy grew richer. This reduction is similar among formal and all workers in the economy and when comparing agriculture to both services and manufacturing. The wage gap decline also coincided with a decline in the share of overall workers employed in agriculture — from 24 percent to 14 percent among all workers and from 19 to 10 percent among workers aged 18 to 65 in the sample. Moreover, the wage gap decline is not 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 differences in the composition of these variables between sectors did not drive the decline during this period. Collectively, the evidence suggests that both the gap level and the decline are 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 differences in mean wages. 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 gap level can indeed be generated by this model. In a second stage of analysis, I use the model to explore productivity growth, a compression of skill differences, and skill-biased technological change as potential drivers of the wage gap decline. I find that all of these can generate a large decline in the wage gap and a significant exit of workers out of agriculture.

The rest of the paper is structured as follows. Section 2 provides a literature review that relates this paper to the literature on wage and output per worker gaps between sectors. Section 3 describes the datasets used. Section 4 describes the magnitude and evolution of wage 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

Recent studies have assessed the role of worker characteristics in explaining inter-sectoral wage gaps across countries.² Vollrath (2014) finds that large wage differences exist between workers after controlling for observed human capital in a set of 14 countries. He finds a limited role of distortions preventing workers from being paid the value of their marginal product in each sector. Similarly, using a different sample of countries, Herrendorf and Schoellman (2015) regress wages on observables allowing for returns on observables to vary

²Additionally, there have been several studies on US labor markets exploring the role of worker characteristics in explaining sectoral wage gaps. These include Krueger and Summers (1988), Murphy and Topel (1987; 1990), and Gibbons and Katz (1992).

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 datasets. 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 migrations³, by Herrendorf and Teixeira (2017) to study sectoral gaps in the US, and by Hicks et al. (2017) to study sectoral wage gaps in 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

³Other studies on migration include Beegle, Weerdt and Dercon (2011), Bryan, Chowdhury and Mobarak (2014), Chiquiar and Hanson (2005), and Yang (2006).

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. 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 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 across production units (Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Herrendorf and Teixeira (2011)). If present at the sectoral level, distortions generating sectoral productivity gaps can be related to the agricultural wage gap—particularly in the presence of large mobility frictions. In the context of Brazil, labor mobility distortions have been used to explain the effects of reducing trade barriers (Dix-Carneiro and Kovak (2014), Dix-Carneiro and Kovak (2017), Adao (2016)) and adopting new technologies (Bustos, Caprettini and Ponticelli (2016)) on local labor markets.

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 a set of Brazilian household surveys from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD) from 1996 to 2013.⁴ This contains a representative sample of all households in Brazil. The survey includes both formal and informal workers and records demographic and employment-status characteristics as well as measures of earnings and hours. Monthly earnings in PNAD include all income in cash for the main job of the reporting week⁵ and monthly hours are constructed from reported weekly hours worked in the main job. Hourly wages are defined as the ratio of these two measures. Worker demographic characteristics include gender, race⁶, age, and educational attainment. 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.

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 Geografia e Estatística* (IBGE)—value added per worker for each year and sector. Sectors are classified into agriculture, manufacturing (including industrial activities) and services.⁷

⁴Both PNAD and RAIS data are available for this period. Although earlier years are available for a large subset of Brazilian workers in RAIS, the lack of universal coverage in earlier periods can be particularly problematic when studying transitions out of agriculture. Hence, the analysis is restricted to this later period. For PNAD, we use all available representative surveys from the period which exclude 2000, 2003, and 2010.

⁵Using other income measures in PNAD such as total income from main occupation and total household monthly income from all household members produces similar results.

⁶Race categories include white, black, brown, asian, and other.

⁷Constructed based on activity or line of business reported. Service categories include construction, trade services, auxiliary services, transport and communication, social services, and public administration.

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, in specifications controlling for differences in unobservable characteristics, administrative formal sector data is used.

Administrative 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.⁸ 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.⁹

The RAIS dataset reports average monthly gross labor earnings including regular salary 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, and these are standardized to match the PNAD measures.

Finally, to identify the employment sector and occupation of workers, classification is

⁸It 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.

⁹IDs available are anonymized to protect the identity of both workers and firms.

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. Similar to the classification used in PNAD, the subsectors are aggregated in to the categories of agriculture, manufacturing¹⁰, 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 of the RAIS 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 using either PNAD or RAIS data, I restrict the analysis to workers between 18 and 65 years old with contracted or reported hours of at least 30 hours a week. Table 1 provides key summary statistics for all workers (PNAD's formal and informal workers) and formal workers (RAIS) data for three sub-periods: 1996-2001, 2002-2007, and 2008-2013. Some features of the data are worth noting. First, the number of formal and informal workers from PNAD (labeled "All") is much lower than the ten percent sample from the RAIS data. This is due to the survey nature of the PNAD and the full coverage of formal workers from the administrative RAIS dataset. Second, mean wages are lower in the PNAD data than in the RAIS data. Other than the inclusion of informal workers in the PNAD, there are several measurement discrepancies between PNAD and RAIS that influence this difference. These include the fact that PNAD wages constructed from cash income reported by workers, which is a noisy measure of after-tax labor earnings. In contrast, the RAIS measure is pre-tax, reported by employers, and includes bonuses and non-cash compensation. To the extent that these discrepancies are larger in agriculture than in other sectors, this would explain part of the difference in the mean agricultural wage gaps measured in the PNAD and RAIS

¹⁰Includes mining industries.

samples. In this paper, models are estimated using the PNAD and RAIS separately to avoid comparing different measures in the same model. The patterns of the agricultural wage gap in both samples are further discussed in the next section. Unlike wages, means of age and education across sectors are similar in both databases, with PNAD workers being only slightly older and less educated than formal workers in RAIS.

When comparing demographics between workers in agriculture and workers in other sectors, education is lower in agriculture in Brazil relative to other sectors in both the PNAD and RAIS samples. During the period of study, educational attainment improved substantially in both samples across all sectors, 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 over time. The explanatory power of age and education will be one of the focal points of the analysis in the following sections. Finally, although wages between agriculture and other sectors are quite different, there are only small gaps 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 of 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 wages between non-agriculture and agriculture among all workers (both formal and informal)¹¹ in the economy declined from 2.6 in 1996 to 1.8 by 2013. Earnings gaps among the same group declined from 2.4 to 1.7. As discussed above, some estimations require the panel structure of the data so that workers can be followed over time — a feature that is only available for formal workers. Among formal workers, the patterns look similar, with the agricultural wage gap declining from 2.3 in 1996 to 1.6 in 2013, and the earnings gap declining from 2.3 to 1.7 (Figure 1). Moreover, both for all and formal workers, the magnitude of the gap and its decline has

¹¹Formal and informal workers in the sample of workers aged 18 to 65 working over 30 hours a week.

Table 1: RAIS SUMMARY STATISTICS

Period	Sector	(1) # Worker-years	log(Wages)						Education		Age				
			(2)		(3)		(4)		(5)		(6)		(7)		
			All	Formal	All	Formal	All	Formal	All	Formal	All	Formal	All	Formal	
1996-2001	Agriculture	0.08	0.97	0.71	1.25	0.84	0.61	1.04	1.10	0.28	0.42	38.18	34.36	13.27	11.21
	Manufacturing	0.08	3.53	1.72	1.87	0.87	0.84	1.43	1.43	0.80	0.74	33.77	32.33	10.78	9.86
	Services	0.37	12.62	1.65	1.90	0.93	0.86	1.59	1.69	0.95	0.93	35.41	34.41	11.28	10.67
	All	0.54	17.12	1.53	1.86	0.97	0.86	1.47	1.60	0.87	0.89	35.63	33.98	11.65	10.57
2002-2007	Agriculture	0.09	1.28	0.83	1.40	0.87	0.52	1.07	1.15	0.33	0.46	38.51	34.49	13.12	11.07
	Manufacturing	0.11	4.41	1.65	1.85	0.81	0.77	1.60	1.62	0.84	0.78	34.70	32.60	11.14	10.01
	Services	0.47	16.34	1.63	1.87	0.87	0.79	1.77	1.90	1.00	0.96	36.15	34.70	11.45	10.88
	All	0.67	22.03	1.53	1.84	0.90	0.78	1.63	1.80	0.94	0.92	36.25	34.27	11.71	10.76
2008-2013	Agriculture	0.07	1.34	1.16	1.69	0.89	0.49	1.15	1.28	0.47	0.57	40.34	35.70	12.79	11.22
	Manufacturing	0.10	5.53	1.87	2.04	0.73	0.71	1.76	1.81	0.90	0.80	35.54	33.34	11.39	10.37
	Services	0.50	22.76	1.87	2.04	0.79	0.73	1.93	2.04	1.05	0.94	37.18	35.04	11.72	11.18
	All	0.68	29.63	1.80	2.02	0.82	0.72	1.81	1.96	1.01	0.91	37.28	34.75	11.86	11.05

Note: Workers aged 18 to 65 for which wages are reported. Formal workers are a 10 percent sample from RAIS. All workers refer to both formal and informal workers reported in PNAD surveys. Number of worker-years are in millions. Wages refer to average monthly earnings divided by hours in real terms (Using 2014 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.

been similar when comparing agriculture to both services and manufacturing individually. In contrast to the differences between agriculture and non-agriculture, mean earnings and wages in the two non-agricultural sectors were similar throughout this period.¹²

Both the levels and decline of the wage gap are significant when compared to other estimates in the literature. For instance, Brazil's 1996 wage gap is above the median of 2.0 from the 12 country sample in Herrendorf and Schoellman (2015). By 2013, Brazil's wage gap falls below this median. Compared with the list of countries from Vollrath (2014), Brazil's 1996 gap between agriculture and manufacturing would rank second highest. When comparing agriculture vs services, the rank would be fourth, just above Indonesia. In contrast, Brazil's 2013 gap levels with respect to manufacturing and services would rank fifth and eleventh, respectively. Although the data on Brazil is not entirely comparable to the wage data from other countries, the significant move down the ranking of countries suggests that Brazil's decline cannot be described as an insignificant change. Moreover, paralleling the results from the output per worker gap literature, Brazil also experienced significant levels and declines in value added per worker gaps between 1996 and 2013.¹³

In tandem to the closing of both output per worker and wage gaps, Brazil also endured a substantial transformation of the employment structure. The economy employed 24 percent of the labor force in agriculture in 1996, which declined to 14 percent by 2013.¹⁴ Manufacturing employed 13 to 15 percent of workers throughout this same period, and services increased from 62 to 72 percent. When analyzing the shares among employed workers aged 18 to 65 in the PNAD sample, agricultural employment declined from 19 to ten percent, manufacturing employment oscillated around 15 to 17 percent, and service employment increased from 65 to 75 percent. Among formal workers in the RAIS sample, where agricultural workers are underrepresented, a similar qualitative pattern is observed with the share of workers in agriculture declining from five to four percent, in manufacturing from 22 to 18 percent, and rising in services from 72 to 78 percent. Altogether, I conclude that the sizeable agricultural wage

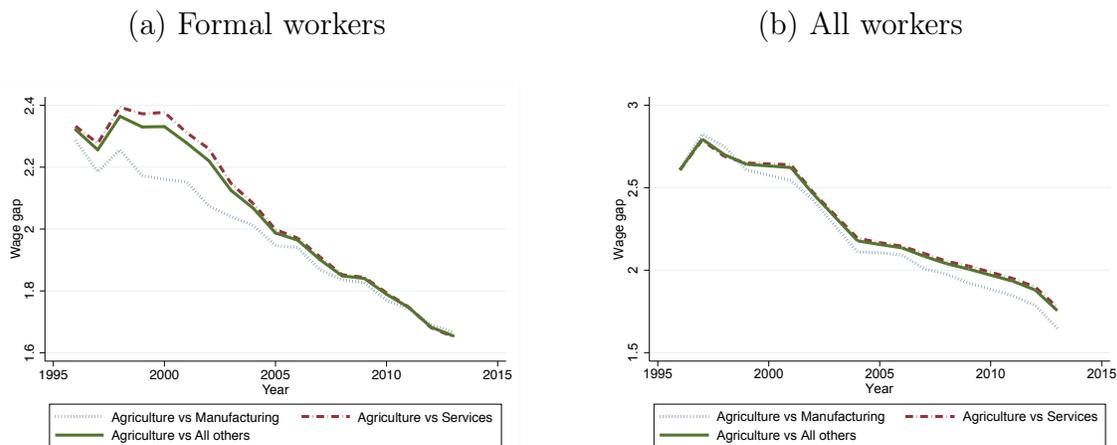
¹²These characteristics are also present when comparing percentiles instead of means. The gaps declined across all percentiles of the wage distribution, with the exception of the poorest tenth percentile of workers in the PNAD. These patterns are documented in appendix A. See Alvarez et al. (2017) and Engbom and Moser (2017) for a more detailed description of the overall inequality compression.

¹³These patterns are discussed in appendix B.

¹⁴Total labor force shares including workers aged 10 and above. The evolution of these shares for the PNAD sample and RAIS are included in appendix C.

gap, its decline, and the shifting of labor away from agriculture and into services are features present among both administrative data on formal workers and survey data including both formal and informal workers. 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.

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.

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 human capital and other worker observables. 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 Worker observables and human capital

To assess the role of observable compositional differences, I first regress wages on a vector of worker observables and sector dummies for every year in the sample. Using both formal and informal workers, the following model is estimated:

$$\begin{aligned} \log(w_{it}) &= \alpha_m^t * M_{it}\phi_t + \alpha_s^t * S_{it}\phi_t + \sum_{a,e} \alpha_{ae}^t * \mathbf{1}(age_{it} = a, edu_{it} = e) \\ &+ \alpha_g^t * gender_{it} + \alpha_r^t * race_{it} + \phi_{state}^t + \epsilon_{it} \end{aligned}$$

where M_{it} and S_{it} are indicators for working in the Manufacturing and Services sectors respectively, age_{it} is age in years, edu_{it} is educational attainment, $gender_{it}$ is a gender indicator, $race_{it}$ is a vector of race indicators and ϕ_{state}^t are year varying state dummies.¹⁵ Coefficients on sector-time interactions, α_s^t and α_m^t , reflect the average wage gap¹⁶ between agriculture and both manufacturing and services in each year t after controlling for observable compositional differences.

Figure 2 depicts the evolution of α_m^t and α_s^t over time using different sets of controls. The baseline wage gap estimates without controls are shown in solid blue. These reflect the difference in means of log wages between non-agriculture and agriculture. Adding a fully interacted set of age and education dummies (dashed red line) reduces the manufacturing-agriculture wage gap by 15 log points (15 percent) in 1996 and by 17 log points (25 percent) in 2013. Similarly, the services-agriculture wage gap is reduced by 26 log points (28 percent) in 1996 and by 26 log points (39 percent) in 2013. Interestingly, adding gender and race as controls (dashed green) fail to reduce the agricultural gap further. This is true in every year of the sample, which indicate that compositional differences in gender and race are not behind average wage differences between agriculture and other sectors. Finally, with

¹⁵All coefficients are allowed to vary over time; therefore, the specification above is equivalent to running cross-sectional regressions for every year.

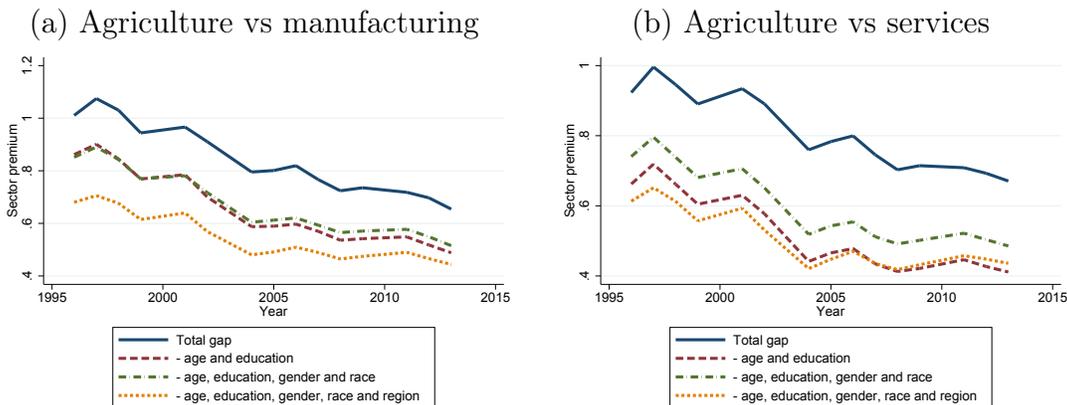
¹⁶For the rest of the paper, I will define the wage gap as the mean difference of log hourly wages with respect to agriculture. Specifically, the gap between sector s' and agriculture is defined as

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

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 studied.

respect to the age and education specification, adding state dummies reduces the gap by an additional 13–17 log points (14–17 percent) in 1996 and 5–7 log points (7–11 percent) in 2013. Overall, worker observables and regional controls explain an average of 35–38 percent of the total wage gap level, of which human capital —as measured by age and education— explain an average of 23–36 percent.

Figure 2: RESIDUAL GAP AFTER CONTROLLING FOR OBSERVABLES



Note: Total gap refers to the difference in mean log wages between sectors for formal and informal workers (PNAD). The lower three lines refer to the residual difference after controlling for differential composition of worker observables (fully interacted vector of age and education, gender, race and states).

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. Table 1 indeed shows differences in education between sectors, with agricultural workers being on average less educated than their peers in services and manufacturing. On the other hand, even if the *composition* of human capital is the same in the two sectors, the *returns*¹⁷ to human capital might be different in the two sectors. To make this distinction, I first estimate the following model for each sector and year.

$$\log(w_{ist}) = F_{st}(\text{education}_{ist}, \text{age}_{ist}) + \epsilon_{ist}$$

Similarly as before, to impose minimal restrictions on how age and education influence wages, the mapping of education and age to wages is specified as $F_{st}(\text{education}_{ist}, \text{age}_{ist}) =$

¹⁷Return differences refer to differences in pay between sectors for each age-education group.

$\beta_t^s X_{it}^s$, where X_{it}^s is a vector of dummies for each age-education group in sector s . Thus, the specification allows full flexibility in terms of age and education, and this relationship can vary in every sector and year. These estimates are then used to conduct a Oaxaca decomposition with agricultural workers as the reference group (Oaxaca (1973)). The wage gap in each year can be decomposed into three components:

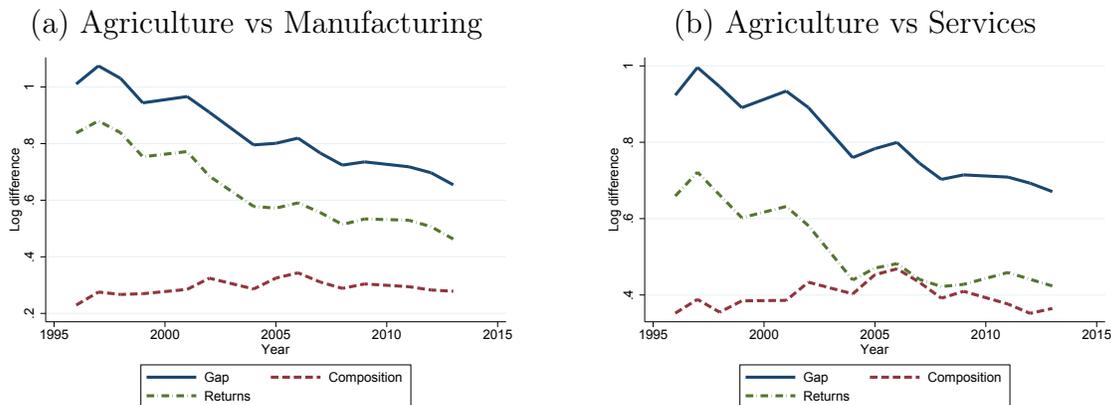
$$\begin{aligned} \Delta_{s'}(E(\log(w_{ist}))) &= \beta_t^a (E(X_{it}^{s'}) - E(X_{it}^a)) + (\beta_t^{s'} - \beta_t^a) E(X_{it}^a) \\ &+ (E(X_{it}^{s'}) - E(X_{it}^a)) (\beta_t^{s'} - \beta_t^a) \end{aligned}$$

The first term is entirely due to composition effects due to age and education differences in workers employed by sector s' relative to 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 in agriculture. 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 composition and return effects.

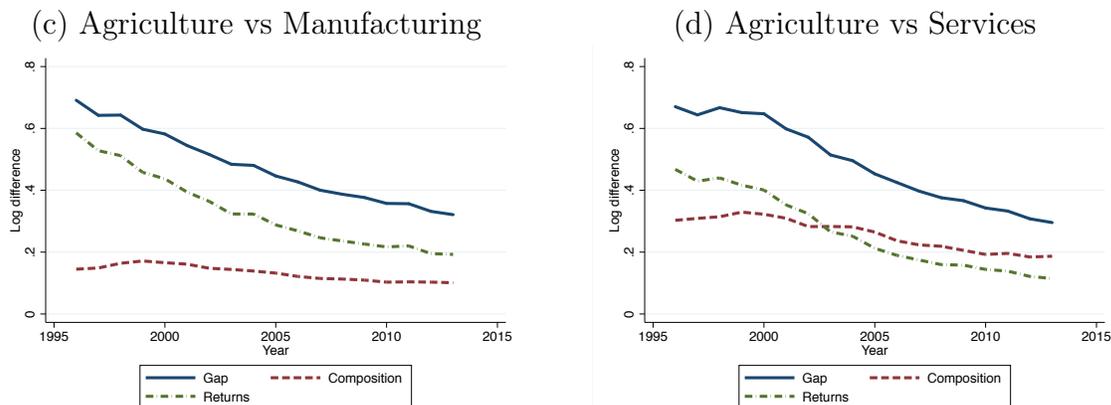
Figure 3 shows the result of this decomposition. Composition effects explain only a small share of the agriculture vs manufacturing gap throughout the sample period, and they explain a larger share, but not all, of the services vs agriculture gap. Differences in age and education cannot account for most 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, virtually all of the decline in the gap between agriculture and both manufacturing and services is driven by the steeper decline in the return component. The pattern is similar when using both survey data on formal and informal workers (PNAD), as well as administrative data from formal workers (RAIS).

Figure 3: OAXACA DECOMPOSITION

ALL WORKERS (PNAD)



FORMAL WORKERS (RAIS)



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

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. To make this distinction empirically and assess the role of unobservables, cross-sectional data is insufficient. For this reason, I focus on the

RAIS panel dataset on formal workers for the remainder of the empirical analysis.

The two competing explanations —differences in unobservables vs differences in pay— have different implications for the behavior of sector-switchers. 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 and predict that sector-switches should be associated with large gains in compensation.

In order to distinguish differential pay from compositional differences in unobservable 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; and ϕ_t and ϕ_i^p are time and individual fixed effects.¹⁸ Individual fixed effects are allowed to vary by six-year periods, but are fixed within each period p .¹⁹ This is done to allow for long-term changes in the distribution of unobservable characteristics. Sector indicators are interacted with time; therefore, the coefficients β_s^t and β_m^t reflect average wage changes from switching out of agriculture into 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

¹⁸Since 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.

¹⁹There are three periods in the sample: 1996–2001, 2002–2007, and 2008–2013.

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,²⁰ and controls are estimated using information from all formal workers in the data.

The time series of both services premiums (β_m^t) and manufacturing premiums (β_s^t) are shown in Figure 4. 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 the period. 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 E). The modest magnitude of premiums 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

²⁰The number of switchers by type of transitions in each year is shown in appendix D.

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 F), it is not entirely clear whether the wage jump from exiting agriculture is drastically different than 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, suggesting that above average long-term gains from transitioning are not large.

Finally, I have also explored whether there are larger gains from workers switching out of the agricultural sector while moving from rural to urban areas at the same time. To the extent that average wage gaps are driven by mobility costs from moving from rural to urban areas, these should be reflected in large gains from moving out of rural agriculture into a new sector at an urban area. The following specification is estimated:

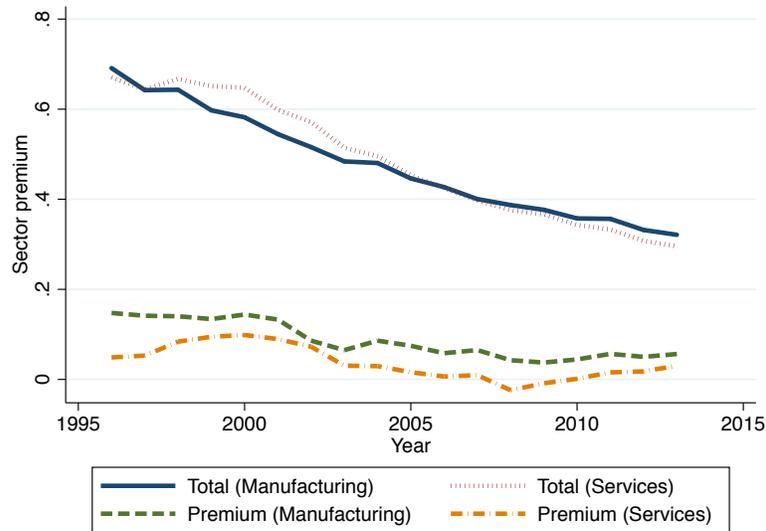
$$\begin{aligned} \log(w_{it}) = & \eta_m^t * M_{it}\phi_t + \eta_s^t * S_{it}\phi_t + \delta_m^t * city_{it} * M_{it}\phi_t + \delta_s^t * city_{it} * S_{it}\phi_t \\ & + \lambda^t * city_{it} * \phi_t + \phi_t + \phi_i^p + \varepsilon_{it} \end{aligned}$$

where $city_{it}$ is an indicator for working in a city. A city is here defined as a municipality with more than one hundred thousand formal workers in 1996. There are 38 out of 5,570 municipalities²¹ that classify as cities under this definition, and they are home to around half of the workers in the sample. This specification allows for the decomposition of sector premiums into the ones estimated from sector transitions alone, η_s^t and η_m^t , and the sector

²¹To overcome problems from changing boundaries over time, I use time consistent municipality definitions (áreas mínimas comparáveis) from IPEA (Reis, Pimentel and Alvarenga (2007)).

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 5. 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 results show significant but moderate additional compensation gains from switching into urban areas on top of switching sectors. These gains are still much smaller than overall wage gap magnitudes.

Figure 4: SECTOR GAPS RELATIVE TO AGRICULTURE CONTROLLING FOR INDIVIDUAL FIXED EFFECTS



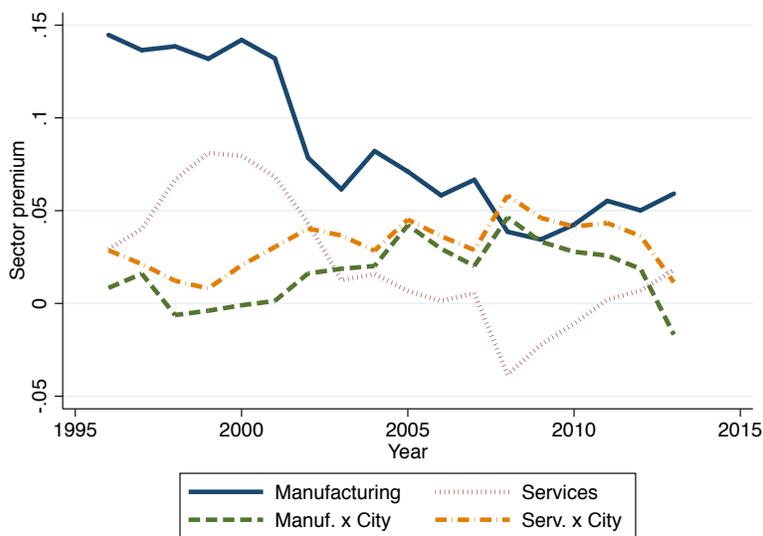
Note: Total refers to the difference in mean log wages between each non-agricultural sector and agriculture. Sector premiums for services (β_s^t) and manufacturing (β_m^t) are defined by equation (1). With the exception of services in 2010, coefficients are all statistically different from zero ($p 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

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.

Figure 5: SECTOR PREMIUMS WITH TRANSITIONS INTO CITIES



Note: Sector premiums for services (η_s^t), manufacturing (η_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.

6 How are compositional differences sustained in equilibrium?

The analysis above suggests that most of the wage gap level is not due to differential pay of equally skilled workers between agriculture and non-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.²² 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. Finally, I explore potential drivers 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

²²See Lagakos and Waugh (2013) and Young (2013).

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

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

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, $\phi_{occupation}^p$ are occupation fixed effects²³ at the three-digit classification level, and the rest of variables are defined as described

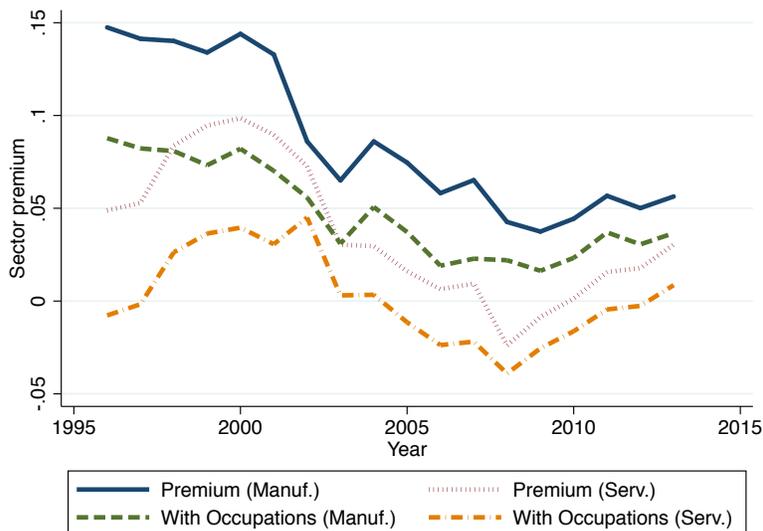
²³Occupation fixed effects are allowed to vary by six-year periods but are fixed within the period.

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_m^t and γ_s^t are identified using sector-switchers that do not switch occupations after they transition.²⁴ In the data, several occupations are common to all sectors (e.g. cleaning, security services, drivers/messengers) and the model is therefore identified.

Figure 6 shows the evolution of premiums with (β_m^t, β_s^t) and without (γ_m^t, γ_s^t) 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.

²⁴The sector premiums estimated in section 5.2 can be written as $\beta_{s'}^t = \gamma_{s'}^t + (E_{s'}(\phi_{occupation}) - E_{agriculture}(\phi_{occupation}))$ where s is each non-agricultural sector and $E_{s'}(\phi_{occupation})$ is the average of occupation fixed effects in each sector.

Figure 6: 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 skill-biased technological change. 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 decreases as income grows.²⁵ 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 positive support. A worker can freely choose to work in one of the two sectors but faces a cost if he decides to work in the non-agriculture sector.²⁶ This distortionary mobility friction is motivated by the compensation differences for workers that switch sectors observed in the data. In this environment, workers maximize their labor income, so that $y_i^l = \max\{w_a(z_a^i), (1 - k)w_n(z_n^i)\}$, where $w_a(z_a^i)$ and $w_n(z_n^i)$ are the wages offered to worker i in the two sectors. In order to focus on the role of friction-induced misallocation and avoid other aggregate income effects stemming from the magnitude of k , I assume that all workers receive a transfer t equivalent to the mobility costs paid in the economy.

²⁵See 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.

²⁶Since 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.

6.2.2 Technology

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

$$Y_a = AZ_a, Y_n = AZ_n$$

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 $Z_s = \int_{i \in \Gamma^s} z_{is}^{\gamma_s} dG$, 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 $L_s = \int_{i \in \Gamma^s} dG$. 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

$$Ap_a z_{ia}^{\gamma_a} \geq Az_{in}^{\gamma_n} (1 - k)$$

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, consumption, given income²⁷ y^i , can be written as follows.

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

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

$$\begin{aligned} \int c_n^i dG &= \int_{i \in \Gamma^n} Az_{in}^{\gamma_n} dG \\ \int c_a^i dG &= \int_{i \in \Gamma^a} Az_{ia}^{\gamma_a} dG \\ \Gamma^n &= \{i : Ap_a z_{ia}^{\gamma_a} \leq Az_{in}^{\gamma_n} (1 - k)\} \\ \Gamma^a &= \{Ap_a z_{ia}^{\gamma_a} > Az_{in}^{\gamma_n} (1 - k)\} \end{aligned}$$

In the model, the wage gap in terms of mean log wages can be written as

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

²⁷Income is the sum of labor income and the transfer, $y^i = \max\{Ap_a z_{ia}^{\gamma_a}, Az_{in}^{\gamma_n} (1 - k)\} + t$.

Differences in mean wages are 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 declines²⁸, 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 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 yearly 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.

²⁸As 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 more detailed discussion.

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 agriculture²⁹ of 19 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 ($\bar{a} = 22.6$) is 22 percent of the average wage and $\phi = 277$.

Production and friction parameters The technology parameter A is set to one hundred 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.³⁰

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 0.3 percent. This reflects the conclusion from the previous section that 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. Since I do not observe all workers at all

²⁹This refers to the share of agricultural workers aged 18 to 65 in the PNAD sample.

³⁰Alternatively, 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 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.

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,

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

The sector-specific distributions have dispersion parameters θ_a and θ_n , which control the within sector variance of the productivity distribution in agriculture and non-agriculture. These are calibrated to match the standard deviation of log wages in agriculture and non-agriculture of 0.84 and 0.94 respectively.³¹ The calibrated parameters that match these log variances are $\theta_a = 2.0$ and $\theta_n = 1.3$.

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.³² Second, the shape 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

³¹From workers aged 18 to 65 in the PNAD sample.

³²By 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.

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 (0.6 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 ($Ap_a z_a^i = Az_n^i(1 - k)$). 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 1.4, which implies a linear correlation of 83 percent between the log of 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 ratio of mean wages between non-agriculture and agriculture is 2.8 in the model. This is seven percent higher than the 2.6 wage gap observed in the PNAD sample. Importantly, the magnitude of the gap is not the result of the mobility friction k . Re-estimating the model with $k = 0$ yields a slightly lower wage gap of 2.6, which is closer to the 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 7) as well as predicted agricultural labor shares (Figure 8).

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 7–8. Growth in the productivity A parameter generates a fall in the wage gap from 2.8 to 2.0, compared with the fall from 2.6 to 1.8 observed in the data (solid orange line). In addition, the increase in aggregate productivity reduces the share of workers employed in agriculture from 19 to eight percent, compared with the decline from 19 to 10 percent observed in the data. Similar results are obtained if one introduces growth exclusively in the agricultural sector.³³

Skill-biased technological change Technology adoption has been proposed as a driving force behind wage dynamics and the exit of workers out of agriculture.³⁴ In Brazil, for instance, 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 within the agricultural sector. To explore this type of skill-

³³Higher productivity growth in the agricultural sector reduces demand for agricultural goods and agricultural workers. This depresses wages in agriculture and produces an exit of workers out of agriculture, resulting in similar wage gap dynamics.

³⁴Ngai 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.

biased technological change, I 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, I 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 0.95 and γ_n decreases from one to 0.78 from 1996 to 2013. The lower latter values indicate a flatter mapping of skill differences to marginal products and, therefore, to wages.

The dashed-dotted green lines in Figures 7–8 show the predicted wage gaps and agricultural labor shares, respectively. Skill-biased technological change induces a decline in the wage gap from 2.8 to 1.1 and a decline in the share of agricultural employment from 19 to 13 percent. Thus, skill-biased technological change in the model generates a larger decline of the wage gap and a smaller decline of the agricultural employment share when compared to the aggregate growth exercise.

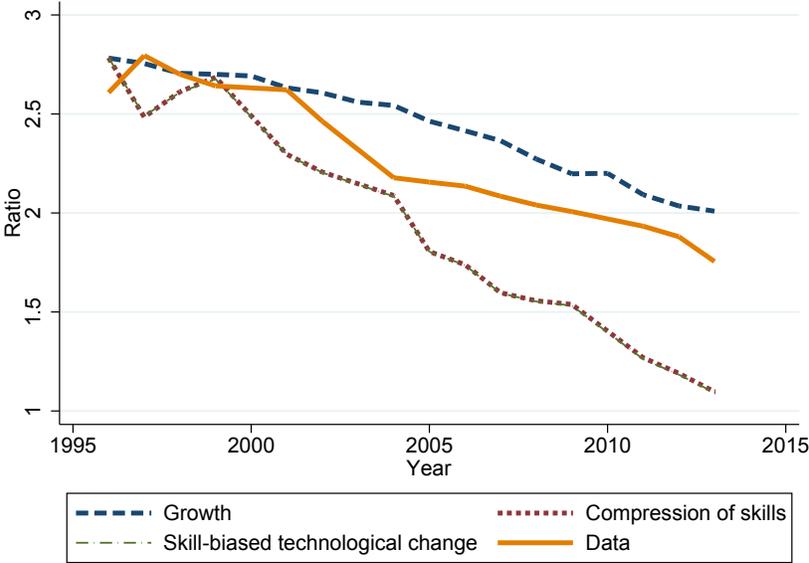
The results suggest that skill-based technological change can explain part of a decline in the agricultural wage gap. However, changes in the mapping of skills to marginal products are at odds with commonly proposed stories of skill-biased technological change in agriculture. In particular, the model 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 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 in the model. 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 variation. This results in an increase in θ_a from 2.0 to 2.1. and an increase in θ_n from 1.3 to 1.7. The dashed red line in Figure 7 shows the gaps generated by the compression in the distribution of worker skills. The results are very similar to the ones induced by skill-biased technological change, with a reduction in the gap from 2.8 to 1.1 and a reduction of agricultural employment from 19 to 13 percent (Figure 8).

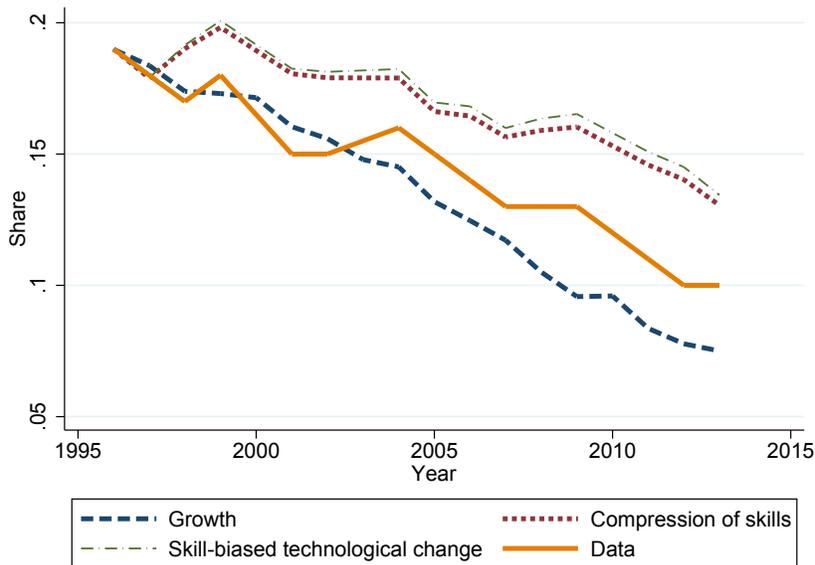
Comparing the three drivers of the wage gap studied, aggregate growth predicts a stronger exit out of agriculture and only a slightly more moderate decline in the wage gap relative to the data. In contrast, a compression in the distribution of marginal products of labor –due either to skill-biased technical change or a compression of skills– exaggerate the wage gap decline while predicting a more moderate exit of workers out of agriculture. Given the simple and abstract nature of the model and calibration, small deviations from projections made over the 17 year period should not be overinterpreted. Altogether, the results show that, in an environment where wage gaps are driven by sorting, growth, skill-biased technological change and a compression of the skill distributions are all consistent with both a significant exit of workers out of agriculture and a reduction of the aggregate agricultural wage gap.

Figure 7: AGRICULTURAL WAGE GAPS



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

Figure 8: SHARE OF WORKERS IN AGRICULTURE



7 Conclusion

The large wage gaps between agriculture and other sectors appear to be driven by compositional differences in worker characteristics. Consistent with cross-country patterns in 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 structural change model of worker sorting is consistent with the magnitude of the aggregate wage gaps observed, the small sectoral premiums estimated in the micro-data, the exit of workers out of agriculture, and the decline of the wage gap over time.

The results highlight the importance of understanding the nature of unobserved worker heterogeneity, since differences in educational attainment and demographics fail to explain most of the aggregate agricultural wage gap. Understanding unobserved differences might also uncover the ultimate sources of the wage gap decline, which remain unclear. On the one hand, the skill-based technological change needed to reduce gaps require low-skill workers to be relatively more valued over time when compared to high-skill ones. This is not easily reconcilable with common narratives of agricultural mechanization and intermediate input usage. On the other hand, little of the compression of the skill distribution can be explained

by changes in observed worker characteristics, including the improvements in educational attainment observed in Brazil during the period studied. Thus, what sources of heterogeneity drive the wage gap level and its compression remains an open question. Unobserved differences in cognitive development, educational quality, and access to job networks are promising alternatives that are consistent with the evidence and might be fruitful avenues for future research. Future wage gap models, however, should be consistent with the main empirical findings presented. They must produce large differences in average pay between sectors without producing 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.
- Adao, Rodrigo. 2016. “Worker Heterogeneity, Wage Inequality, and International Trade.”
- Alvarez, Jorge, Felipe Benguria, Niklas Engbom and Chris Moser. 2017. “Firms and the Decline in Labor Earnings Inequality in Brazil.” *American Economic Journal: Macroeconomics* (forthcoming) .
- 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 Weerd 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).

- Dix-Carneiro, Rafael and Brian Kovak. 2014. "Trade Liberalization and Labor Market Dynamics." *American Economic Review (forthcoming)* 82(3).
- Dix-Carneiro, Rafael and Brian Kovak. 2017. "Trade Liberalization and Regional Dynamics." *American Economic Review (forthcoming)* .
- Donovan, Kevin. 2016. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences."
- Engbom, Niklas and Chris Moser. 2017. "Earnings Inequality and the Minimum Wage: Evidence from Brazil." *Working Paper* .
- 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 Arilton Teixeira. 2017. "Wages, Human Capital, and Structural Transformation."
- 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.
- Reis, Eustaquio, Marcia Pimentel and Ana Isabel Alvarenga. 2007. “Areas Minimas Comparáveis Para os Períodos Intercensitários de 1872 a 2000.”. IPEA.
- 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. "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* 128(4):1727–1785.

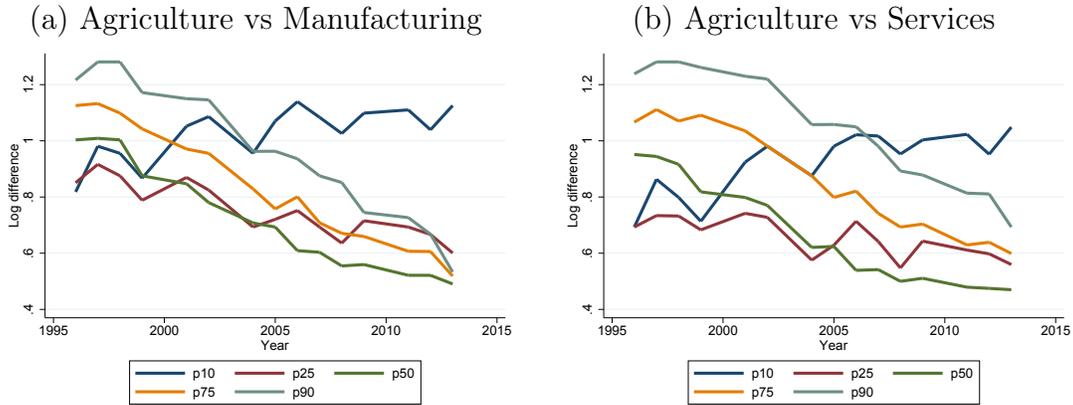
Appendix

A The decline in wage gaps throughout the distribution

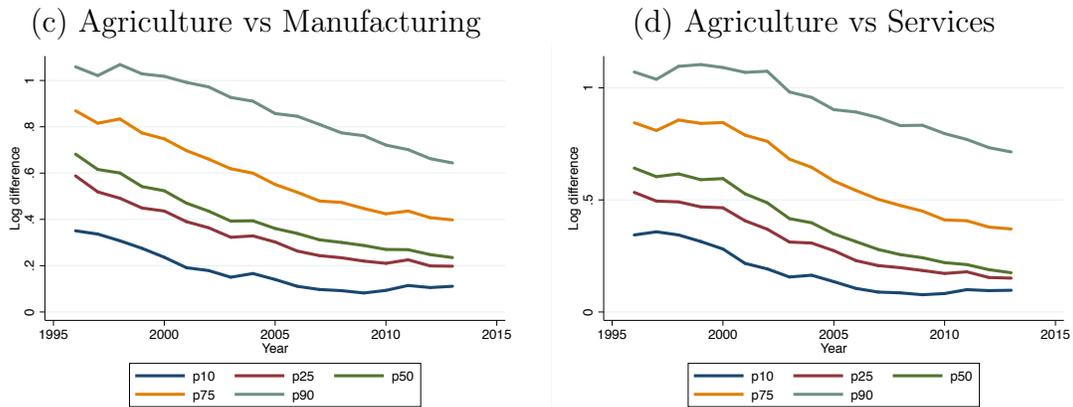
Figure 9 shows the ratio of wage percentiles in agriculture and non-agriculture for both the PNAD and RAIS samples. Percentiles are here defined by the ranking of workers within each sector. The decline in the agricultural wage gap is present throughout the wage distribution in both datasets, with the exception of the lowest tenth percentile of workers in the PNAD. 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, so 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.

Figure 9: GAPS IN BRAZIL BY PERCENTILE

ALL WORKERS (PNAD)



FORMAL WORKERS (RAIS)



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.

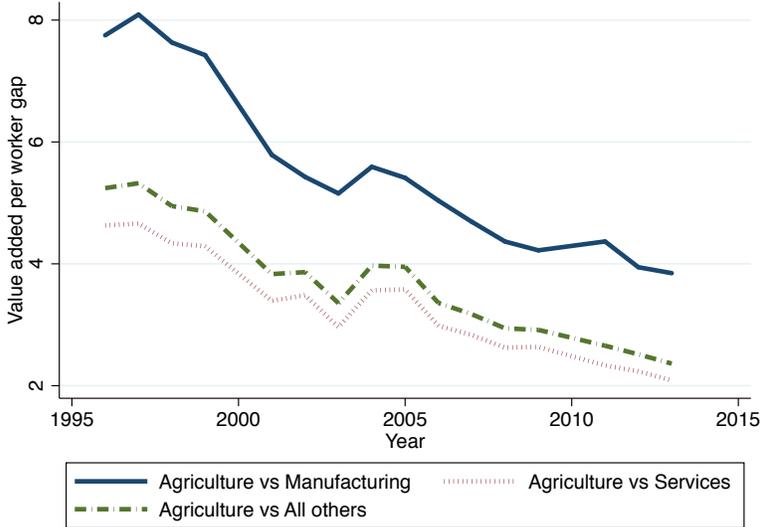
B The evolution of output gaps in Brazil

This section documents the magnitudes and decline of output gaps in Brazil and how both output and wage gap magnitudes compare to estimates from the literature. Figure 10 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 agriculture-manufacturing output gap is much larger than the agriculture-services gap. This is expected due in part to the natural differences in capital intensities between services and agriculture. These differences notwithstanding, the declining pattern of pay and value added per worker gaps is qualitatively similar.

Similarly to wage gaps, 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 24 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.

Figure 10: VALUE ADDED PER WORKER GAP IN BRAZIL

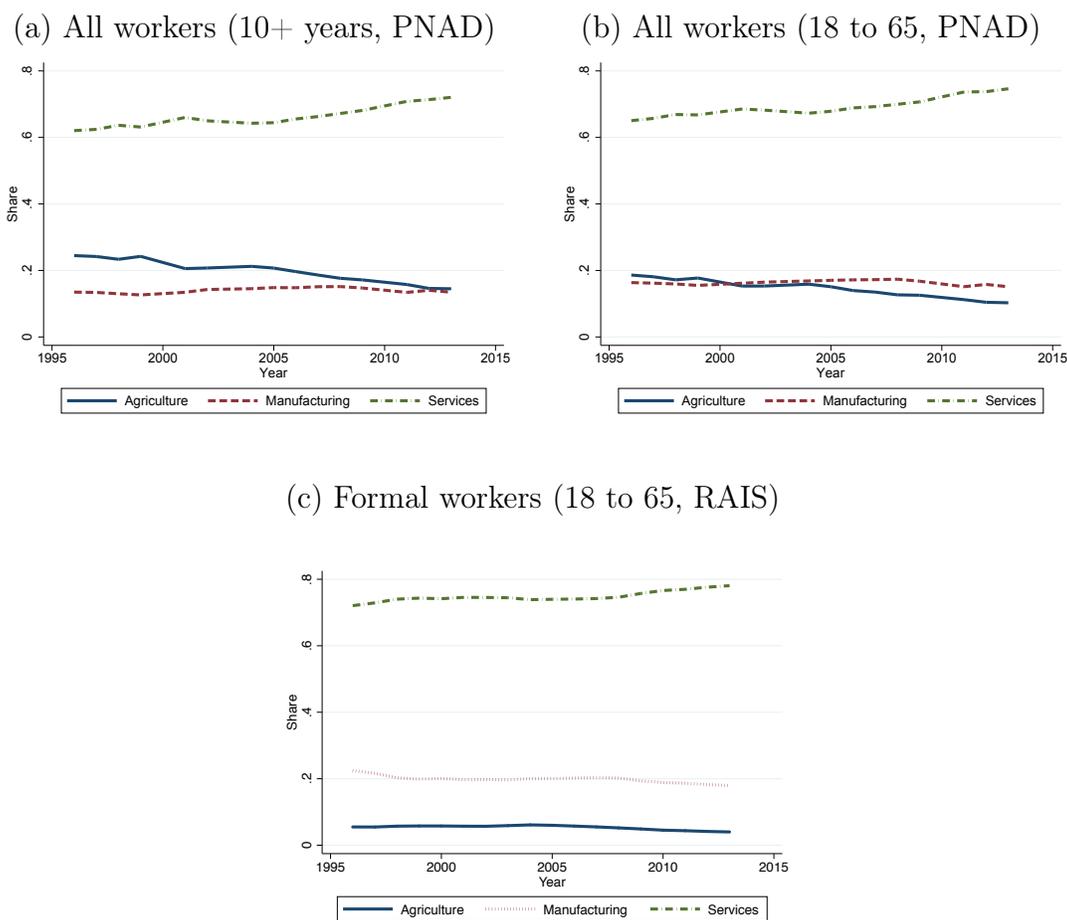


Note: Value added per worker gaps are constructed from national accounts available from IBGE and labor statistics from the PNAD.

C The evolution of agricultural labor shares

Employment shares by sector based are shown in Figure 11. Figure (a) shows data from all workers in the PNAD. Figure (b) shows data from workers in the PNAD sample aged 18 to 65 working more than 30 hours a week. Figure (c) shows the equivalent for formal workers in the RAIS.

Figure 11: WORKERS BY SECTOR



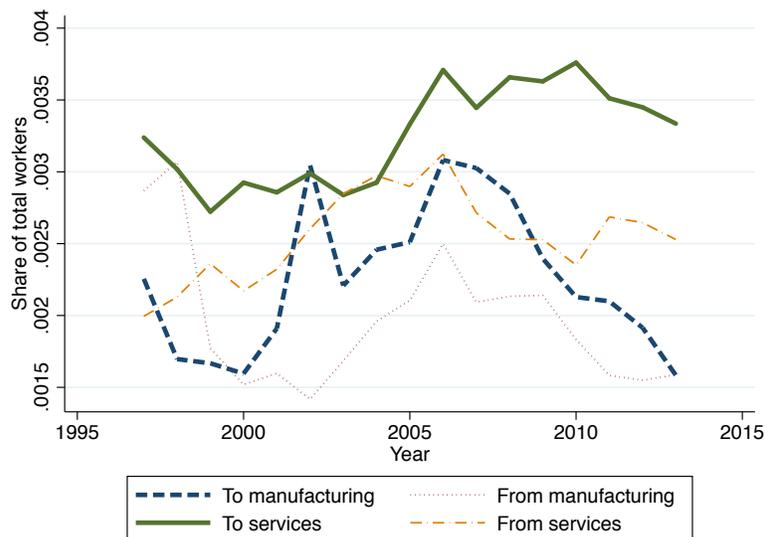
Note: Share of total employed workers. All workers refer to formal and informal workers in PNAD. Formal worker data is from RAIS. Figures (b) and (c) contain workers reporting more than 30 hours of weekly work.

D Sector-switchers over time

Figure 12 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 12: NUMBER OF SECTOR-SWITCHERS FROM AND INTO AGRICULTURE

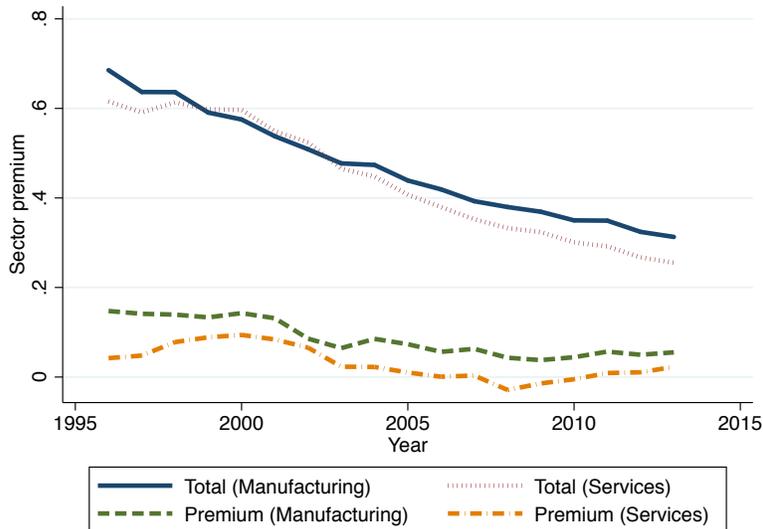


Note: Share of total employed workers that switch out of or into agriculture in any given year.

E Sector premiums using earnings

The analogous results from Figure 4 using monthly earnings instead of hourly wages as a dependent variable are shown in Figure 13. 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 13: 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 agriculture. 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 ($p < .01$).

F Event-study of transitions into and out of agriculture

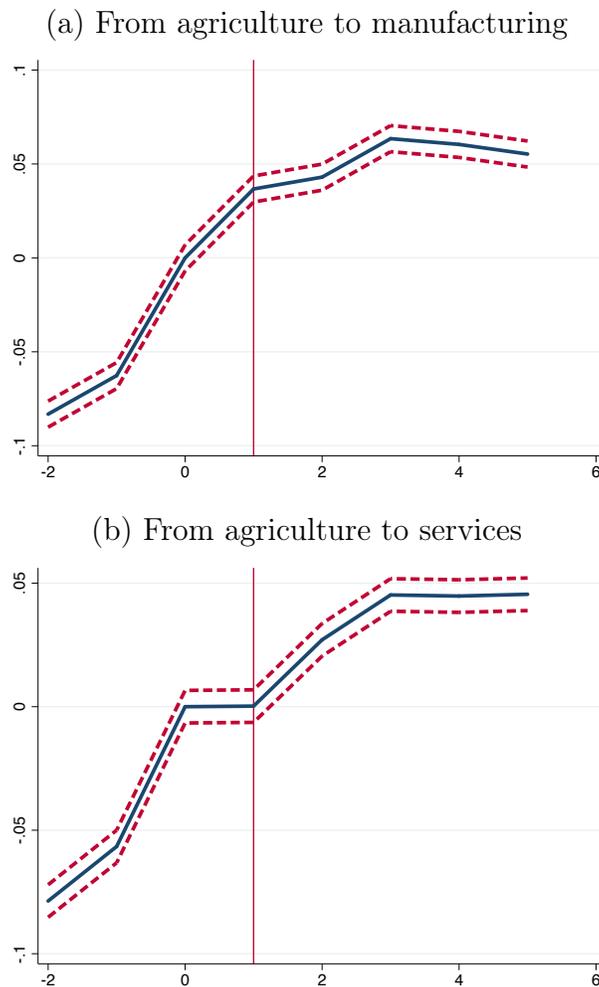
This section adopts an event-study framework focusing on workers that switch out of agriculture. The following equation is estimated for transitions out of agriculture into both services and manufacturing.

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

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.

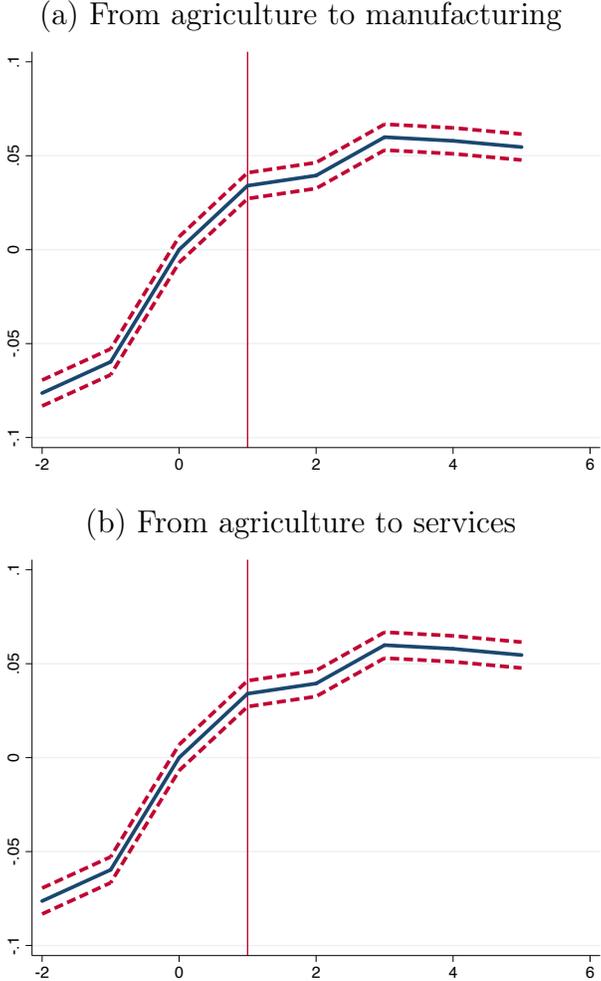
Figure 14 shows the results of the transition coefficients, γ_j , with confidence intervals. Transitions into services and manufacturing are analyzed separately with similar results. As before, the wage increases five years after transitions are much smaller than the magnitude of the aggregate wage gap. 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 15 shows similar results when age squared is added as an additional control.

Figure 14: TRANSITIONS OUT OF AGRICULTURE



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 15: TRANSITIONS OUT OF AGRICULTURE CONTROLLING FOR AGE SQUARED



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 subtracted 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.