

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

By Jorge A. Alvarez

APPENDIX A: SUPPLEMENTARY SUMMARY STATISTICS

Baseline estimations focus on workers aged between 18 and 65 with contracted or reported hours of at least 30 hours a week. In addition, 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.

This section shows additional summary statistics before and after the sample restrictions and for the self-employed as a separate group. Table A1 shows a comparison of summary statistics for all PNAD observations and PNAD observations in the baseline sample of workers aged 18 to 65 with reported hours above 30 per week. In the PNAD, hours and age restrictions changes the share of workers in agriculture (from 23 to 17 percent) and in services (from 64 to 67 percent), while keeping roughly the same share of manufacturing workers (13 percent) in the earliest period. Changes over time are similar in both raw and restricted PNAD sample. In the RAIS, as reported in Table A2, sectoral shares before and after sample selection are roughly unchanged.

In addition, Table A3 shows the prevalence of self-employment across sectors and compares the distribution of wages among the self-employed to that of the baseline sample. Self-employment is significantly less prevalent among formal workers and workers in manufacturing. Across all sectors and periods, the self-employed earn less on average. Nonetheless, average wage gaps between agriculture and non-agriculture are of a similar order of magnitude and show similar declining patterns as those documented in the baseline sample (see Figure B3).

APPENDIX B: WAGE GAPS FOR ALTERNATIVE SAMPLES AND MEASURES

Wage gap magnitudes and trends are broadly similar using alternative available measures. Figure B1 shows the ratio of wages between non-agricultural sectors and agriculture using the following measures of wages and income: main job cash wage, average cash wage across jobs, all income per hour, labor income (including non-cash) in the main job, labor income (including non-cash) across all jobs, and all income. Similarly, Figure B2 shows wage gaps using household income per person measures with similar results.

In addition, Figure B3 documents average wage gaps for different sample selection criteria. Magnitudes and trends are broadly similar when using all observations, workers of all ages working more than 30 hours a week, and the baseline sample of workers aged 18 to 65. When focusing of the more narrow set of self-employed workers, gaps are slightly lower when comparing agriculture and manufacturing, and slightly higher when comparing agriculture and services. The trends in gaps among the self-employment are similar to those in the baseline sample.

Table A1—: SUMMARY STATISTICS FOR ALL WORKERS VS SAMPLE (PNAD)

Sector	(1) # Worker-years	log(Wages)		Education				Age						
		(2)		(3)		(4)		(5)						
		Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.					
All	All	All	All	All	All	All	All	All	All					
<i>1996-2001</i>														
Agr.	160.4	82.1	0.73	0.71	0.90	0.84	1.04	1.04	0.25	0.28	37.20	38.18	17.90	13.27
Manuf.	97.7	81.9	1.65	1.72	0.91	0.87	1.40	1.43	0.78	0.80	33.16	33.77	12.10	10.78
Serv.	512.3	372.1	1.60	1.65	0.99	0.93	1.55	1.59	0.93	0.95	34.96	35.41	12.85	11.28
All	770.4	536.1	1.48	1.53	1.02	0.97	1.41	1.47	0.83	0.87	35.25	35.63	14.16	11.65
<i>2002-2007</i>														
Agr.	176.7	91.9	0.85	0.83	0.94	0.87	1.06	1.07	0.31	0.33	38.96	38.51	17.68	13.12
Manuf.	129.8	106.0	1.58	1.65	0.89	0.81	1.54	1.60	0.82	0.84	35.01	34.70	12.76	11.14
Serv.	623.7	468.1	1.61	1.63	0.93	0.87	1.72	1.77	1.00	1.00	36.02	36.15	12.82	11.45
All	930.2	666.0	1.51	1.53	0.96	0.90	1.56	1.63	0.91	0.94	36.46	36.25	13.99	11.71
<i>2008-2013</i>														
Agr.	137.5	74.0	1.17	1.16	0.98	0.89	1.12	1.15	0.43	0.47	41.78	40.34	17.19	12.79
Manuf.	119.5	101.5	1.83	1.87	0.82	0.73	1.70	1.76	0.89	0.90	36.15	35.54	12.88	11.39
Serv.	640.3	502.8	1.88	1.87	0.88	0.79	1.89	1.93	1.05	1.05	37.29	37.18	12.96	11.72
All	897.3	678.3	1.80	1.80	0.91	0.82	1.74	1.81	1.00	1.01	37.84	37.28	13.82	11.86

Note: All workers refers to both formal and informal in PNAD surveys. Sample refers to workers aged 18 to 65 for which wages are reported working over 30 hours a week. Number of worker-years are in millions. Percentages expressed are shares of total workers estimates using survey weights. Wages refer to average monthly earnings divided by hours in real terms (using 2014 Reals). Education levels are defined as 1= Primary or middle school or no education, 2= high school 3= some college education and 4= college completed.

Table A2—: SUMMARY STATISTICS FOR ALL WORKERS VS SAMPLE (RAIS)

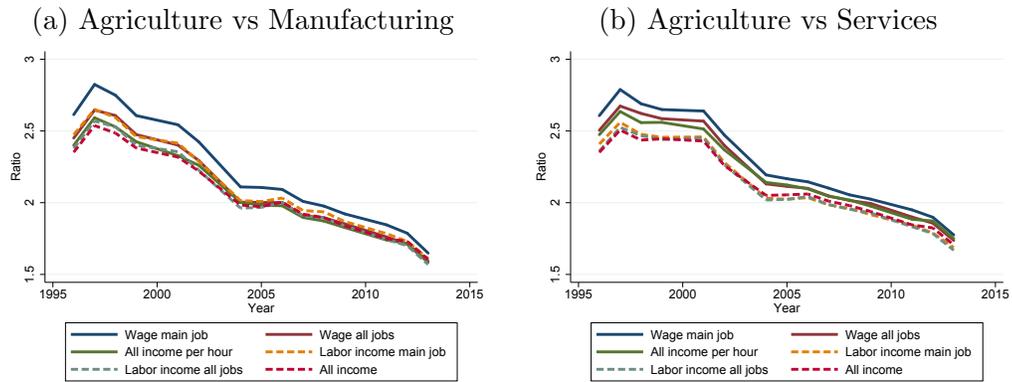
Sector	log(Wages)										Age									
	(1)		(2)				(3)				(4)		(5)		(6)		(7)			
	All	Samp.	# Worker-years	Mean	All	Samp.	Std. dev.	All	Samp.	Mean	All	Samp.	Mean	All	Samp.	Mean	All	Samp.	Std. dev.	Samp.
<i>1996-2001</i>																				
Agr.	1182.6	977.8	1.22	1.25	0.61	0.61	0.61	1.10	1.10	1.10	0.42	0.42	0.42	33.58	34.35	11.53	11.21			
Manuf.	4103.4	3543.6	1.83	1.87	0.84	0.84	1.43	1.43	1.43	1.43	0.75	0.74	0.74	31.76	32.32	10.04	9.86			
Serv.	15585.7	12662.1	1.91	1.90	0.89	0.86	1.73	1.69	1.69	1.69	0.97	0.94	0.94	33.97	34.40	10.82	10.66			
All	20871.7	17183.5	1.86	1.86	0.88	0.86	1.64	1.60	1.60	1.60	0.93	0.89	0.89	33.51	33.97	10.75	10.57			
<i>2002-2007</i>																				
Agr.	1584.4	1290.7	1.37	1.40	0.53	0.52	1.15	1.15	1.15	1.15	0.45	0.46	0.46	34.13	34.48	11.29	11.07			
Manuf.	5138.9	4424.8	1.82	1.85	0.77	0.77	1.62	1.62	1.62	1.62	0.78	0.78	0.78	32.20	32.59	10.13	10.01			
Serv.	20651.1	16390.0	1.89	1.87	0.83	0.79	1.96	1.96	1.96	1.90	1.00	0.96	0.96	34.48	34.69	11.08	10.88			
All	27374.4	22105.5	1.85	1.84	0.81	0.78	1.85	1.80	1.80	1.80	0.96	0.92	0.92	34.03	34.26	10.96	10.76			
<i>2008-2013</i>																				
Agr.	1719.1	1379.2	1.67	1.70	0.49	0.49	1.27	1.28	1.28	1.28	0.56	0.57	0.57	35.34	35.71	11.44	11.22			
Manuf.	6929.9	5770.2	2.01	2.05	0.71	0.71	1.81	1.81	1.81	1.81	0.79	0.80	0.80	32.84	33.37	10.47	10.37			
Serv.	30909.3	23534.0	2.05	2.04	0.77	0.74	2.10	2.04	2.04	2.04	0.97	0.94	0.94	34.74	35.05	11.36	11.17			
All	39558.3	30683.4	2.03	2.03	0.75	0.73	2.01	1.97	1.97	1.97	0.95	0.92	0.92	34.43	34.76	11.24	11.05			

Note: All workers refers to both formal and informal in PNAD surveys. Sample refers to workers aged 18 to 65 for which wages are reported working over 30 hours a week. Number of worker-years are in millions. Percentages expressed are shares of total workers estimates using survey weights. 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.

Table A3—: SELF-EMPLOYMENT IN DIFFERENT SAMPLES

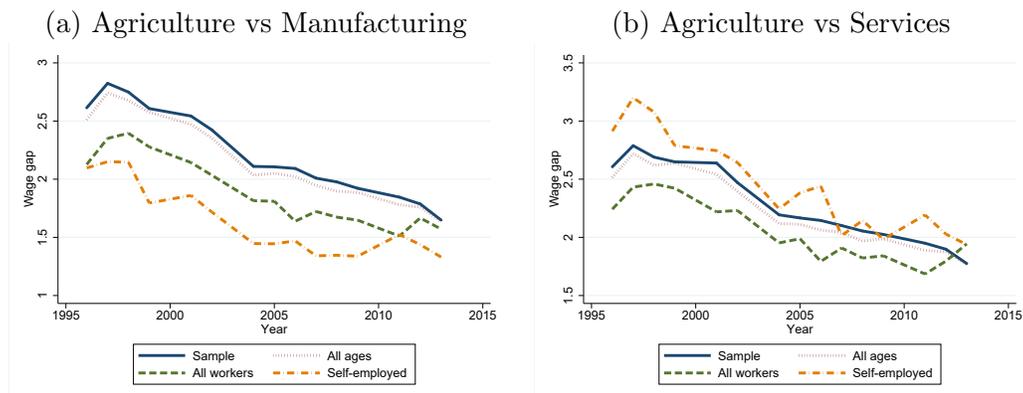
Sector	Share of self-employed (%)						Mean						Std. dev.												
	(1)		(2)		(3)		(4)		(5)		(6)		(1)		(2)		(3)		(4)		(5)		(6)		
	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	PNAD	RAIS	
<i>1996-2001</i>																									
Agr.	26.4	35.7	6.6	7.1	0.71	0.58	1.25	1.03	0.84	0.93	0.61	0.42													
Manuf.	9.7	7.7	1.1	1.1	1.72	1.34	1.87	1.39	0.87	0.94	0.84	0.64													
Serv.	23.9	21.8	2.5	2.5	1.65	1.63	1.90	1.33	0.93	0.95	0.86	0.59													
All	22.6	21.9	2.5	2.5	1.53	1.33	1.86	1.29	0.97	1.05	0.86	0.58													
<i>2002-2007</i>																									
Agr.	25.6	33.3	7.2	7.8	0.83	0.65	1.40	1.24	0.87	1.00	0.52	0.38													
Manuf.	17.5	11.9	1.0	1.0	1.65	1.17	1.85	1.44	0.81	0.92	0.77	0.52													
Serv.	21.3	18.8	2.5	2.6	1.63	1.55	1.87	1.39	0.87	0.93	0.79	0.48													
All	21.6	19.8	2.5	2.6	1.53	1.29	1.84	1.37	0.90	1.02	0.78	0.47													
<i>2008-2013</i>																									
Agr.	27.4	32.6	7.1	7.6	1.16	0.96	1.70	1.53	0.89	1.05	0.49	0.35													
Manuf.	15.9	11.0	1.0	0.9	1.87	1.47	2.05	1.67	0.73	0.87	0.71	0.45													
Serv.	19.9	17.7	2.2	2.3	1.87	1.82	2.04	1.62	0.79	0.86	0.74	0.41													
All	20.5	18.3	2.2	2.3	1.80	1.61	2.03	1.61	0.82	0.97	0.73	0.41													

Note: All workers refers to workers of all ages without any sampling restrictions in each dataset. Sample refers to workers aged 18 to 65 for which wages are reported working over 30 hours a week. Number of worker-years are in thousands. 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.



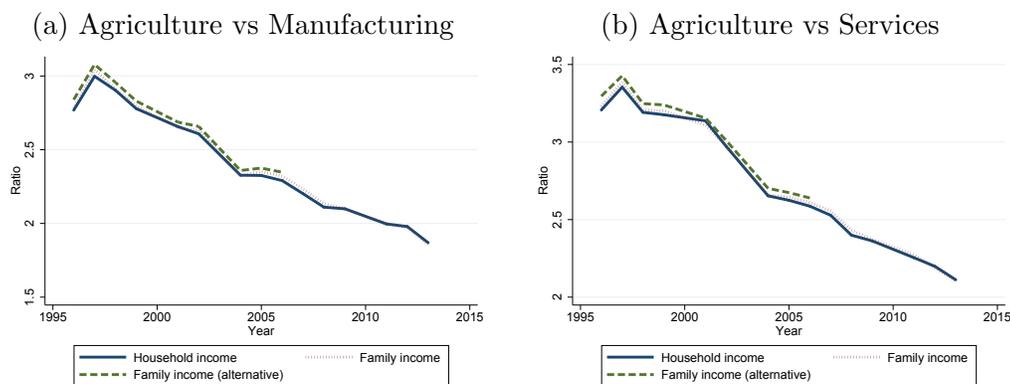
Note: Formal and informal workers from PNAD in baseline selection of workers aged 18 to 65 working a minimum of 30 hours a week. Wages are computed as labor income over reported working hours for both the main job and all jobs. Labor income refers to total income directly related to the main or all occupations. All income refers to income from all sources excluding the income from people whose status in the household was a pension recipient, domestic worker, or relative of an employed domestic worker in the household.

Figure B1. : WAGE GAPS FOR ALTERNATIVE DEFINITION OF WAGES AND INCOME



Note: Formal and informal workers from PNAD. Sample refers to baseline selection of workers aged 18 to 65 working a minimum of 30 hours a week. All ages refer to workers of all ages present in PNAD working a minimum of 30 hours. All workers refer to all workers sampled by PNAD. Wage Gap is the ratio in mean wages between the non-agricultural sector and agriculture.

Figure B3. : WAGE GAPS UNDER ALTERNATIVE SAMPLES



Note: Household from formal and informal workers from PNAD in baseline selection of workers aged 18 to 65 working a minimum of 30 hours a week. Monthly household and family income is presented under three definitions used by PNAD. All of these exclude the income from people whose status in the household was a pension recipient, domestic worker, or relative of an employed domestic worker in the household. Alternative definition only available for the earlier period.

Figure B2. : GAPS IN HOUSEHOLD INCOME PER PERSON

APPENDIX C: THE DECLINE IN WAGE GAPS THROUGHOUT THE DISTRIBUTION

Figure C1 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.

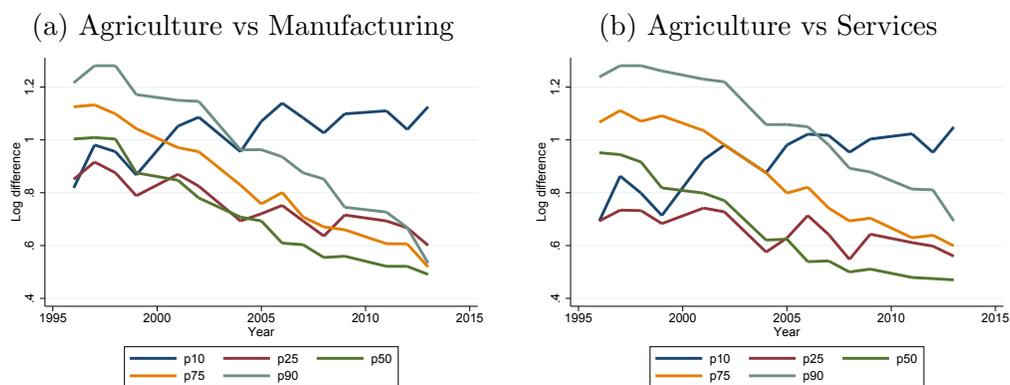
APPENDIX D: THE EVOLUTION OF WAGE AND OUTPUT GAPS IN BRAZIL RELATIVE TO OTHER COUNTRIES

Brazil's 1996 wage gap is above the median of 2.0 from the 12 country sample in Berthold Herrendorf and Todd 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.

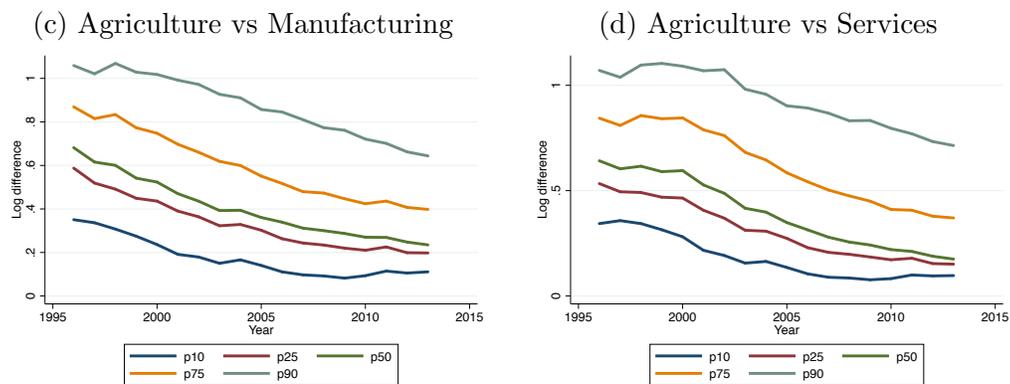
Moreover, paralleling the results from the output per worker gap literature, Brazil also experienced significant levels and declines in value added per worker gaps between 1996 and 2013. Figure D1 shows how the between-sector difference in gross domestic product per worker as measured by the national 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 Berthold Herrendorf and Todd Schoellman (2015) in their 12 country sample and just

ALL WORKERS (PNAD)

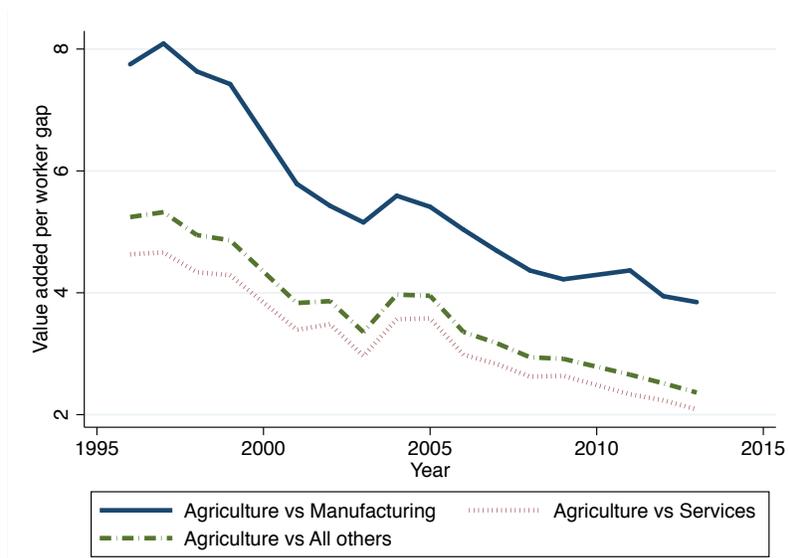


FORMAL WORKERS (RAIS)



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.

Figure C1. : GAPS IN BRAZIL BY PERCENTILE



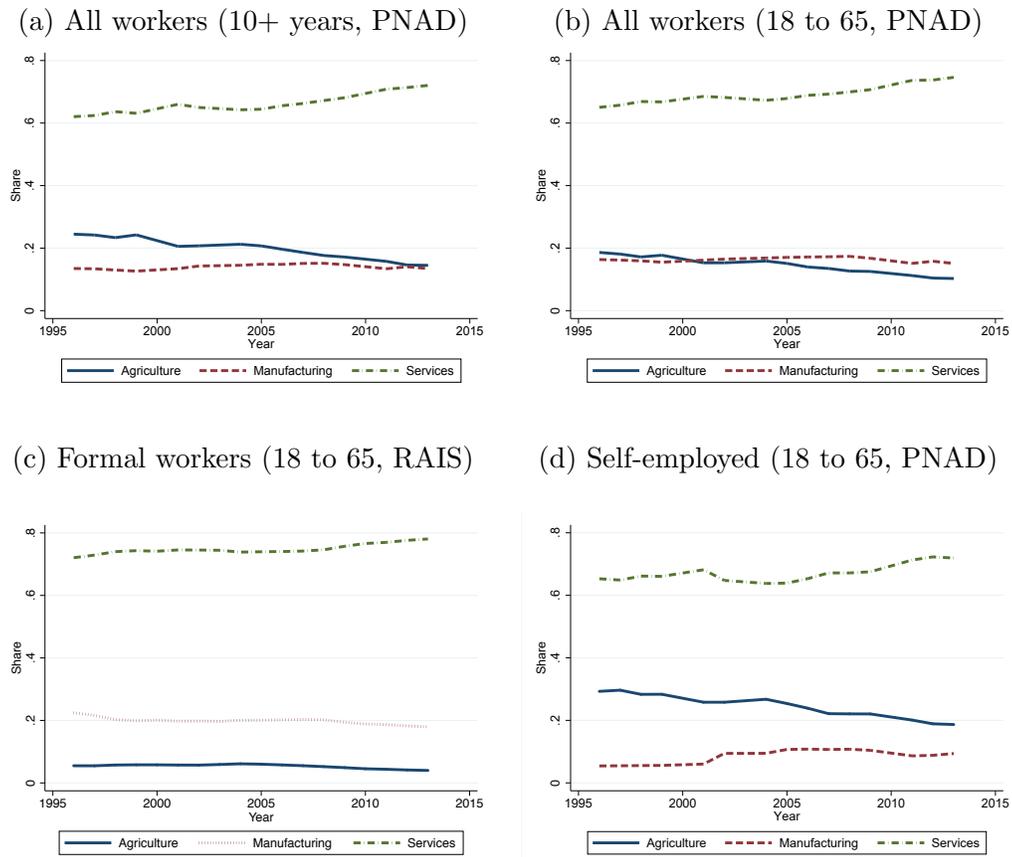
Note: Value added per worker gaps are constructed from national accounts available from IBGE and labor statistics from the PNAD.

Figure D1. : VALUE ADDED PER WORKER GAP IN BRAZIL

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 Berthold Herrendorf and Todd 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.

APPENDIX E: LEVELS AND EVOLUTION OF AGRICULTURAL EMPLOYMENT SHARES UNDER DIFFERENT SAMPLES

Employment shares by sector are shown in Figure E1. 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, and Figure (d) shows employment shares for the self-employed.



Note: Share of total employed workers. All workers refer to formal and informal workers in PNAD without restrictions on working hours. Formal worker data is from RAIS. Figures (b) and (c) contain workers reporting more than 30 hours of weekly work.

Figure E1. : WORKERS BY SECTOR.

APPENDIX F: OAXACA DECOMPOSITION ASSESSING THE ROLE OF OBSERVABLE HUMAN CAPITAL

There are two margins on which observable human capital influences the wage gap. On the one hand, observable human capital can be lower in one sector than the other. On the other hand, even if the *composition* of observable human capital is the same in the two sectors, the *returns*²⁹ 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.

$$\log(w_{sit}) = \mathbf{edu_age}'_{sit} \boldsymbol{\beta}_s^t + \epsilon_{sit}$$

As before, to impose minimal restrictions on how age and education influence wages, $\mathbf{edu_age}_{sit}$ 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:

$$\begin{aligned} \Delta_{\check{s}}(E(\log(w_{\check{s}it}))) &= (E(\mathbf{edu_age}'_{\check{s}it}) - E(\mathbf{edu_age}'_{ait})) \mathbf{b}_a^t \\ &+ E(\mathbf{edu_age}'_{ait}) (\mathbf{b}_{\check{s}}^t - \mathbf{b}_a^t) \\ (F1) \quad &+ (E(\mathbf{edu_age}'_{\check{s}it}) - E(\mathbf{edu_age}'_{ait})) (\mathbf{b}_{\check{s}t} - \mathbf{b}_{at}) \end{aligned}$$

Here, \mathbf{b}_s^t is the vector of estimated coefficients on age and education dummies in year t and sector s . The first term is entirely due to composition differences in age and education between sector \check{s} and agriculture. In other words, this component reflects the mean wage gap if all education-age groups were equally paid in both agriculture and sector \check{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 F1 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

²⁹Return differences refer to differences in pay between sectors for each age-education group.

ALL WORKERS (PNAD)

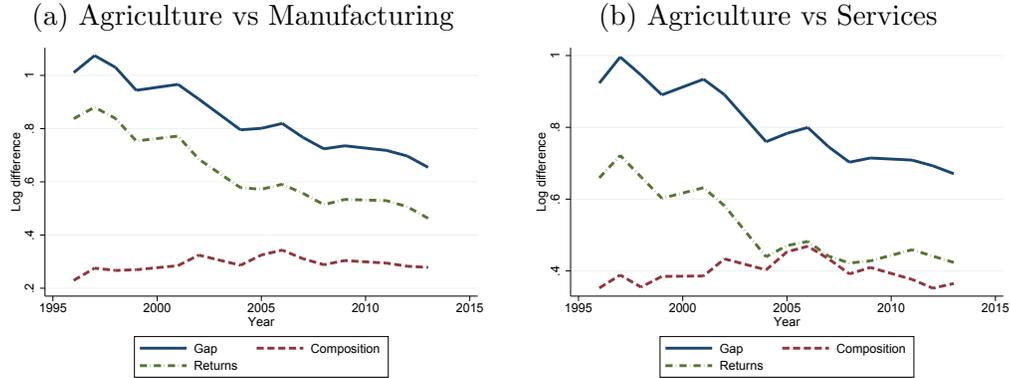


Figure F1. : OAXACA DECOMPOSITION

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).

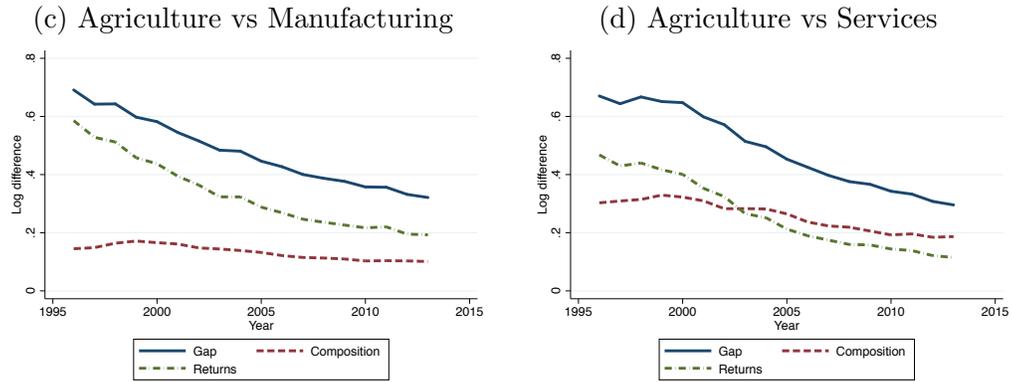
APPENDIX G: SECTOR-SWITCHERS OVER TIME

Figure G1 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.

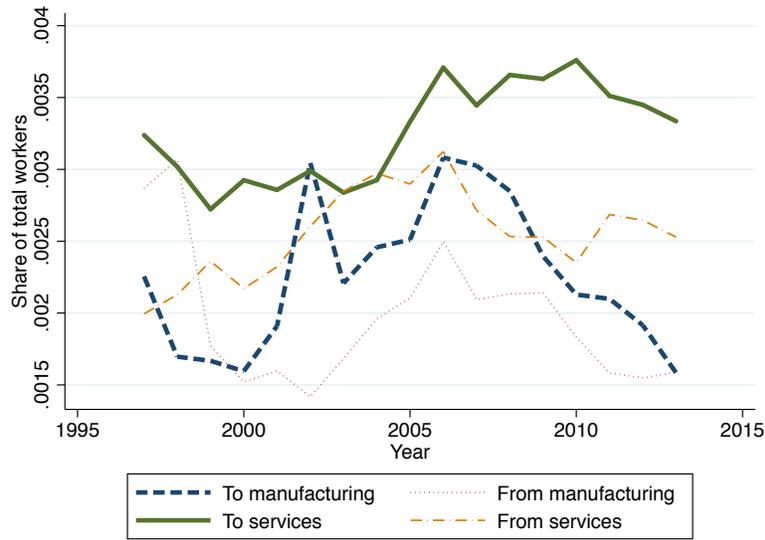
APPENDIX H: SECTOR PREMIUMS USING EARNINGS

The analogous results from Figure 3 using monthly earnings instead of hourly wages as a dependent variable are shown in Figure H1. Similarly to the baseline, earnings premiums are significant but also 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.

FORMAL WORKERS (RAIS)

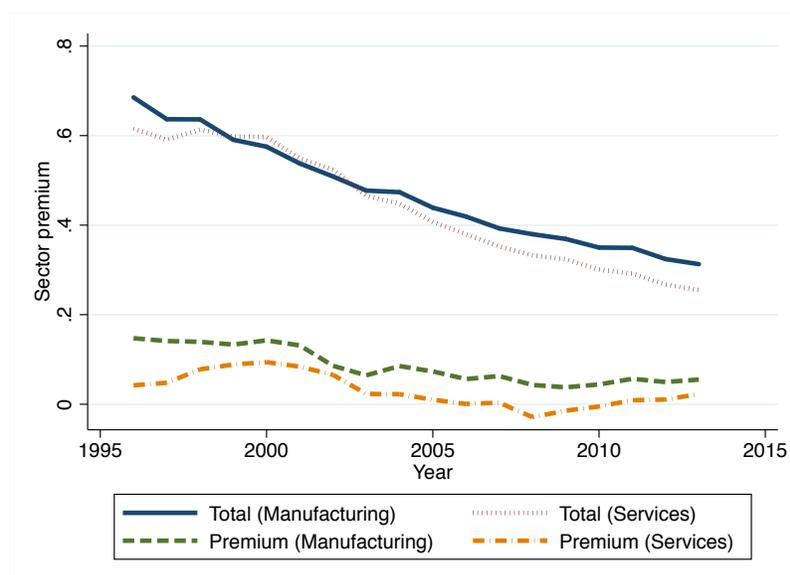


Note: Gap refers to the difference in mean log wages between two sectors. Returns refer to the first term and composition refers to the second term of the Oaxaca decomposition.



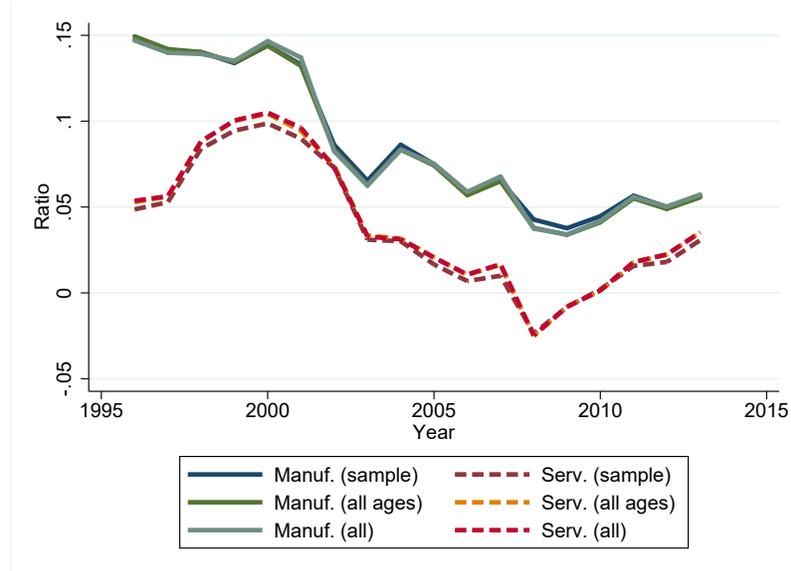
Note: Share of total employed workers that switch out of or into agriculture in any given year.

Figure G1. : NUMBER OF SECTOR-SWITCHERS FROM AND INTO AGRICULTURE



Note: Total refers to the difference in mean log earnings between each non-agricultural sector and agriculture. Sector premiums for services (β_s^T) and manufacturing (β_m^T) are defined by equation (2). With the exception of 2008, coefficients are all statistically different from zero ($p < .01$).

Figure H1. : 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^r) and manufacturing (β_m^r) are defined by equation (2). Sample refers to the baseline estimates. All ages, refers to a selection of workers based only on hours and not ages. All includes all observations in RAIS for which wages are available.

Figure 11. : SECTOR GAPS RELATIVE TO AGRICULTURE CONTROLLING FOR WORKER FIXED EFFECTS UNDER DIFFERENT SAMPLES

APPENDIX I: SECTOR PREMIUMS WITH ALTERNATIVE SAMPLES

Figure 11 shows sectoral premium estimations (controlling for worker fixed effects) using three different samples: all available observations in the RAIS data (including multiple-job records per year), workers of all ages working over 30 hours a week, and the baseline sample of 18 to 65 workers. Results are virtually unchanged across sample specifications.

APPENDIX J: CONTROLLING FOR OCCUPATIONS: EVIDENCE OF SECTOR-SPECIFIC SKILLS

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

Although I cannot test the presence of sector-specific skills directly, I can test an implication of sector-specific skills on wage changes among sector-switchers and the sector premiums described in section IV.B. 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 IV.B are significantly reduced once I control for changes in occupation by estimating

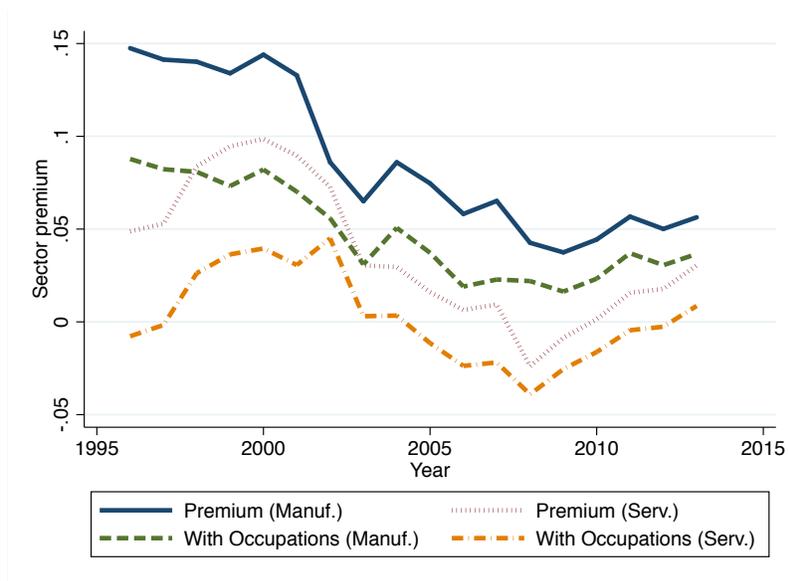
$$(J1) \quad \log(w_{it}) = \sum_{\tau=1996}^{2013} \gamma_m^\tau m_{it} + \sum_{\tau=1996}^{2013} \gamma_s^\tau s_{it} + \phi_{occupation}^p + \phi_t + \phi_i^p + \varepsilon_{it}$$

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 effects³⁰ at the three-digit classification level, and the rest of variables are defined as described in section IV.B. 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^τ and γ_s^τ are identified using sector-switchers that do not switch occupations after they transition.³¹ In the data, several occupations are common to all sectors (e.g. cleaning, security services, drivers/messengers) and the model is therefore identified.

Figure J1 shows the evolution of premiums with $(\beta_m^\tau, \beta_s^\tau)$ and without $(\gamma_m^\tau, \gamma_s^\tau)$

³⁰Occupation categories used are at the three-digit aggregation level. Occupation fixed effects are allowed to vary by six-year periods but are fixed within the period.

³¹The sector premiums estimated in section IV.B can be written as $\beta_{s'}^\tau = \gamma_{s'}^\tau + (E_{s'}(\phi_{occupation}^p) - E_{agriculture}(\phi_{occupation}^p))$ where s is each non-agricultural sector and $E_{s'}(\phi_{occupation}^p)$ is the average of occupation fixed effects in each sector.



Note: Sector premiums for services (β_s^T) and manufacturing (β_m^T) are defined by equation (2). Sector premiums controlling for occupations for services (γ_s^T) and manufacturing (γ_m^T) are defined by equation (5). With the exception of services (without occupations) in 2010, premium coefficients are all statistically different from zero ($p < .01$).

Figure J1. : SECTOR PREMIUMS AFTER CONTROLLING FOR OCCUPATIONAL CHANGES

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.

APPENDIX K: HETEROGENEITY OF SECTOR PREMIUMS

This section presents results of sector premiums by age and education using the RAIS data. The models estimated are of the form

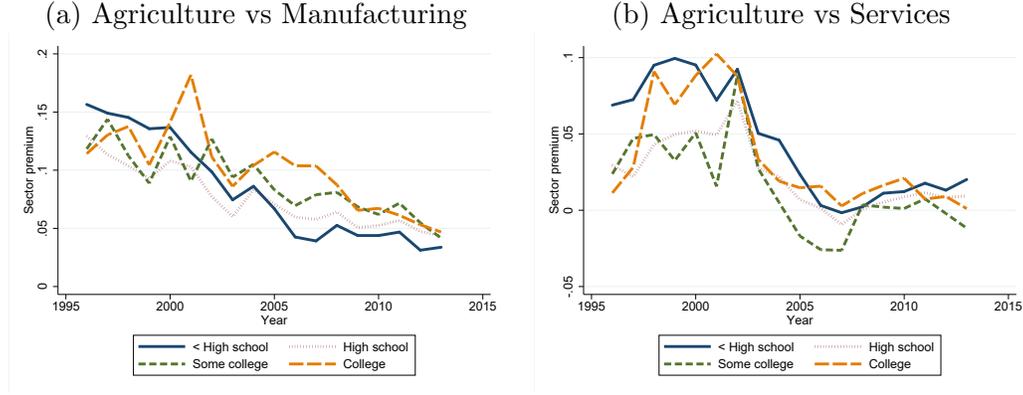


Figure K1. : SECTOR PREMIUMS BY EDUCATION

$$\log(w_{it}) = \sum_{\tau=1996}^{2013} \sum_{e=1}^4 \beta_m^{e\tau} m_{it} \mathbf{1}[edu_{it} = e] + \sum_{\tau=1996}^{2013} \sum_{e=1}^4 \beta_s^{e\tau} s_{it} \mathbf{1}[edu_{it} = e] + \phi_{et} + \phi_i^p + \varepsilon_{it}$$

$$\log(w_{it}) = \sum_{\tau=1996}^{2013} \sum_{a=1}^3 \beta_m^{a\tau} m_{it} \mathbf{1}[agegroup_{it} = a] + \sum_{\tau=1996}^{2013} \sum_{e=1}^4 \beta_s^{e\tau} s_{it} \mathbf{1}[agegroup_{it} = a] + \phi_{at} + \phi_i^p + \varepsilon_{it}$$

where ϕ_{et} and ϕ_{at} are interacted year-education and year-age group dummies respectively. The first model estimates sector premiums for different levels of education (less than high school, high school, some college, and college), while the second model estimates sector premiums for three different age groups (ages 18–30, 31–40, and 40–65). The premiums for manufacturing and services by education and year ($\beta_s^{e\tau}$) are shown in Figure K1, while the premiums by age are shown in Figure K2.

APPENDIX L: ROBUSTNESS OF SECTOR PREMIUMS TO SECTOR-SWITCHER RESTRICTIONS

Table L1 shows the average sector premium coefficients by period when the model in equation (2) is estimated using only sector-switchers and only transitions out of agriculture.

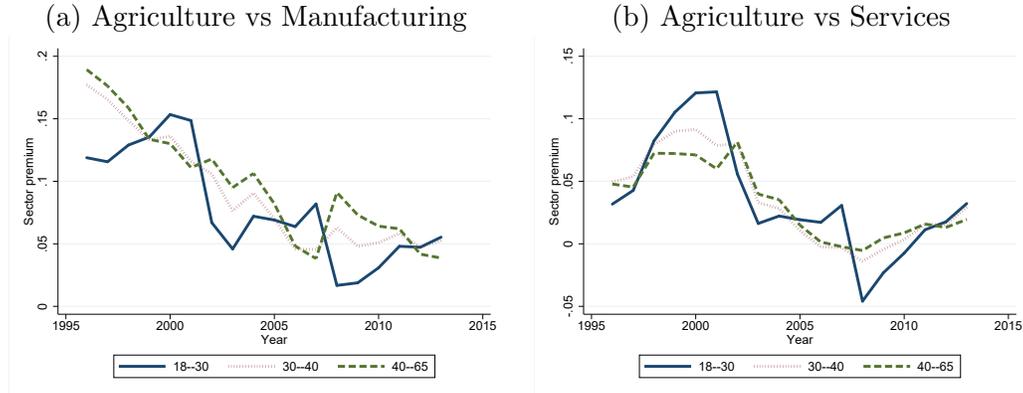


Figure K2. : SECTOR PREMIUMS BY AGE

Table L1—: 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 (2). 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.

APPENDIX M: 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 M1 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 M2 shows similar results when age squared is added as an additional control.

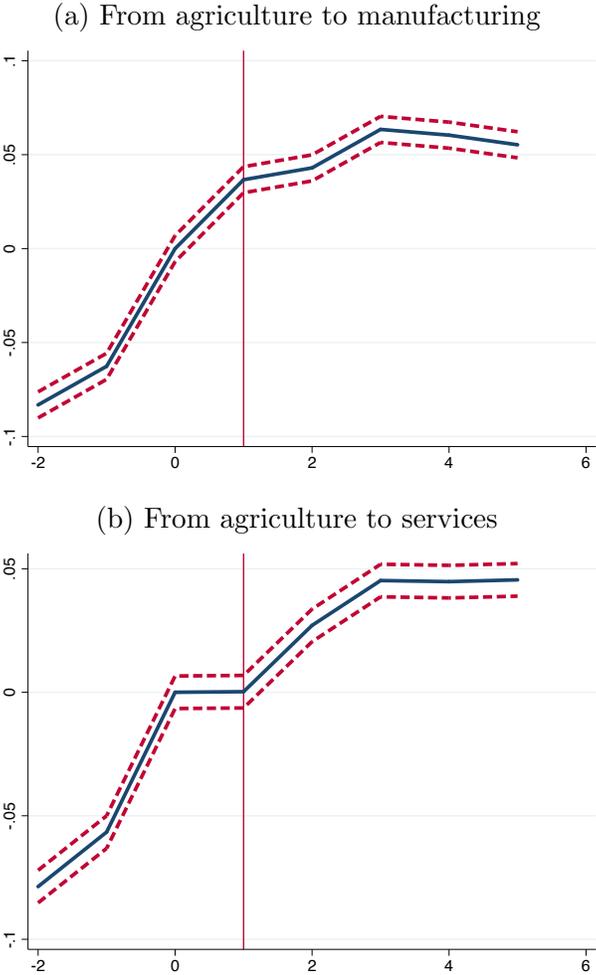
APPENDIX N: TRANSITIONS OUT OF AGRICULTURE INTO CITIES

The following specification is estimated:

$$\begin{aligned} \log(w_{it}) = & \sum_{\tau=1996}^{2013} (\eta_m^\tau m_{it} + \eta_s^\tau s_{it} + \delta_m^\tau m_{it} * city_{it} + \delta_s^\tau s_{it} * city_{it} + \lambda^\tau * city_{it}) \\ & + \phi_t + \phi_i^p + \varepsilon_{it} \end{aligned}$$

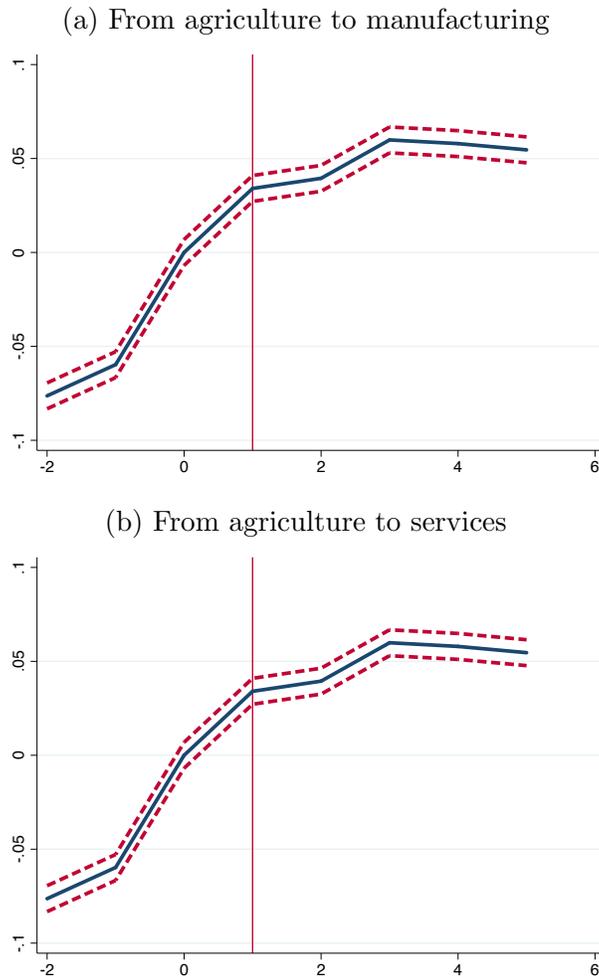
where $city_{it}$ is an indicator for working in a city. A city is here defined as a municipality with more than one hundred thousand formal workers in 1996. There are 38 out of 5,570 municipalities³² that classify as cities under this definition, and they are home to around half of the workers in the sample. This specification allows for the decomposition of sector premiums into the ones estimated from sector transitions alone, η_s^τ and η_m^τ , and the sector premiums associated to transitions between sectors into cities, δ_s^τ and δ_m^τ . The evolution of these coefficients over time is shown in Figure N1. The coefficients on the city and sector

³²To overcome problems from changing boundaries over time, I use time consistent municipality definitions (áreas mínimas comparáveis) from IPEA (Reis, Pimentel and Alvarenga (2007)).



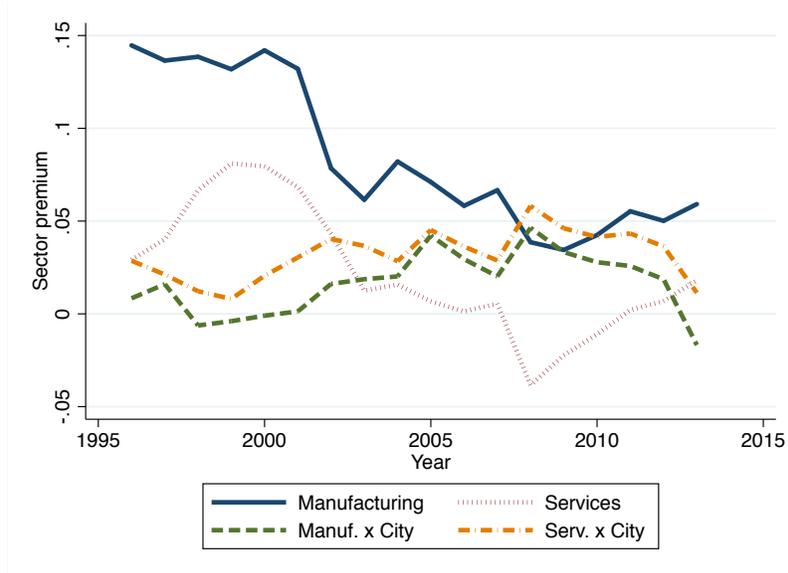
Note: Year 0 refers to the last year worked in the pre-transition sector and Year 1 refers to the first year in the post-transition sector. Solid line shows coefficients γ_j subtracted by γ_0 so that coefficients reflect changes relative to the pre-transition wage level. Dashed lines are 95 percent confidence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

Figure M1. : TRANSITIONS OUT OF AGRICULTURE



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.

Figure M2. : TRANSITIONS OUT OF AGRICULTURE CONTROLLING FOR AGE SQUARED



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

Figure N1. : SECTOR PREMIUMS WITH TRANSITIONS INTO CITIES

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 smaller than overall wage gap magnitudes.

APPENDIX O: CALIBRATION OF THE JOINT DISTRIBUTION OF SECTOR-SPECIFIC SKILLS

In the model, workers have dependent draws from sector-specific Fréchet distributions $X(z_a)$ and $Y(z_n)$. The joint distribution $G(z_a, z_n)$ is a Frank copula resulting from the two primary distributions. That is,

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

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.³³ 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.³⁴ 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 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 12.8, 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.

Since wage gap magnitudes are sensitive to different values of the ρ parameter, robustness exercises are also conducted. Assuming an uncorrelated skill distribution ($\rho = 0$) increases the gap to 3.0, while increasing the correlation to near one or negative one produces wage gaps of 3.7 and 8.7 respectively.³⁵ Although it is hard to estimate accurately the correlation of skills in the population, the results do indicate that large wage gaps of the magnitudes observed in Brazil,

³³From workers aged 18 to 65 in the PNAD sample to include dispersion using formal and informal workers in the economy. Using standard deviations from workers in the RAIS produces an even higher agricultural wage gap.

³⁴By 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.

³⁵The value of ρ that minimizes the wage gap produces a linear correlation of log skills of 67 percent and a wage gap of 2.5.

and their decline, can be generated by compositional differences in skills arising from a sorting equilibrium.