Online Appendices

(Not for publication)

A Data and Definitions

A.1 Tariffs

Tariff data come from Kume et al. (2003), who report nominal tariffs and effective rates of protection from 1987 to 1998 using the Brazilian industry classification *Nível* 50. We aggregate these tariffs slightly to an industry classification that is consistent with the Demographic Census data used to construct local tariff shock measures. The classification is presented in Table A1. In aggregating, we weight each *Nível* 50 industry by its 1990 industry value added, as reported in IBGE National Accounts data. Figure A1 shows the evolution of nominal tariffs from 1987 to 1998 for the ten largest industries. The phases of Brazilian liberalization are visible (see Section 2 for a discussion and citations). Large nominal tariff cuts from 1987-1989 had little effect on protection, due to the presence of substantial nontariff barriers and tariff exemptions. In 1990, the majority of nontariff barriers and tariff exemptions were abolished, being replaced by tariffs providing equivalent protection; note the increase in tariffs in some industries in 1990. During liberalization, from 1990 to 1994, tariffs fell in all industries, then were relatively stable from 1995 onward.

In Section 4.2 we calculate post-liberalization tariff changes using UNCTAD TRAINS. See Appendix B.8 for details.

Table A1: Consistent Industry Classification Across Censuses and Tariff Data

Industry	Industry Name	Nível 50	1970, 1980, 1991 Census (*atividade*)	2000, 2010 Census (*CNAE-Dom*)
1	Agriculture	1	011-037, 041, 042, 581	1101-1118, 1201-1209, 1300, 1401, 1402, 2001, 2002, 5001, 5002
2	Mineral Mining (except combustibles)	2	050, 053-059	12000, 13001, 13002, 14001-14004
3	Petroleum and Gas Extraction and Coal Mining	3	051-052	10000, 11000
4	Nonmetallic Mineral Goods Manufacturing	4	100	26010, 26091, 26092
5	Iron and Steel, Nonferrous, and Other Metal Production and Processing	5-7	110	27001-27003, 28001, 28002
8	Machinery, Equipment, Commercial Installation Manufacturing, and Tractor Manufacturing	8	120	29001
10	Electrical, Electronic, and Communication Equipment and Components Manufacturing	10-11	130	29002, 30000, 31001, 31002, 32000, 33003
12	Automobile, Transportation, and Vehicle Parts Manufacturing	12-13	140	34001-34003, 35010, 35020, 35030, 35090
14	Wood Products, Furniture Manufacturing, and Peat Production	14	150, 151, 160	20000, 36010
15	Paper Manufacturing, Publishing, and Printing	15	170, 290	21001, 21002, 22000
16	Rubber Product Manufacturing	16	180	25010
17	Chemical Product Manufacturing	17,19	200	23010, 23030, 23400, 24010, 24090
18	Petroleum Refining and Petrochemical Manufacturing	18	201, 202, 352, 477	23020
20	Pharmaceutical Products, Perfumes and Detergents Manufacturing	20	210, 220	24020, 24030
21	Plastics Products Manufacturing	21	230	25020
22	Textiles Manufacturing	22	240, 241	17001, 17002
23	Apparel and Apparel Accessories Manufacturing	23	250, 532	18001, 18002
24	Footwear and Leather and Hide Products Manufacturing	24	190, 251	19011, 19012, 19020
25	Food Processing (Coffee, Plant Products, Meat, Dairy, Sugar, Oils, Beverages, and Other)	25-31	260, 261, 270, 280	15010, 15021, 15022, 15030, 15041-15043, 15050, 16000
32	Miscellaneous Other Products Manufacturing	32	300	33001, 33002, 33004, 33005, 36090, 37000
91	Utilities	33	351, 353	40010, 40020, 41000
92	Construction	34	340, 524	45001-45005
93	Wholesale and Retail Trade	35	410-424, 582, 583	50010, 50030, 50040, 50050, 53010, 53020, 53030, 53041, 53042, 53050, 53061-53068, 53070, 53080, 53090, 53101, 53102, 55020
94	Financial Institutions	38	451-453, 585, 612	65000, 66000, 67010, 67020
95	Real Estate and Corporate Services	40, 41	461-464, 543, 552, 571-578, 584, 589	63022, 70001, 71020, 72010, 74011, 74012, 74021, 74022, 74030, 74040, 74050, 74090, 92013, 92014, 92015, 92020
96	Transportation and Communications	36, 37	471-476, 481, 482, 588	60010, 60020, 60031, 60032, 60040, 60091, 60092, 61000, 62000, 63010, 63021, 64010, 64020, 91010
97	Private Services	39, 43	511, 512, 521-523, 525, 531, 533, 541, 542, 544, 545, 551, 577, 586, 587, 613-619, 622-624, 632, 901, 902	1500, 50020, 53111, 53112, 53113, 55010, 55030, 63030, 70002, 71010, 71030, 72020, 73000, 74060, 80011, 80012, 80090, 85011, 85012, 85013, 85020, 85030, 90000, 91020, 91091, 91092, 92011, 92012, 92030, 92040, 93010, 93020, 93030, 93091, 93092, 95000
98	Public Administration	42	354, 610, 611, 621, 631, 711-717, 721-727	75011-75017, 75020

Tradeable (industries 1–32)

Nontradeable (industries 91–98)

Consistent industry classification used in generating local tariff shocks from *Nível* 50 tariff data in Kume et al. (2003) and Decennial Census data.

Figure A1: Tariffs - 1987-1998
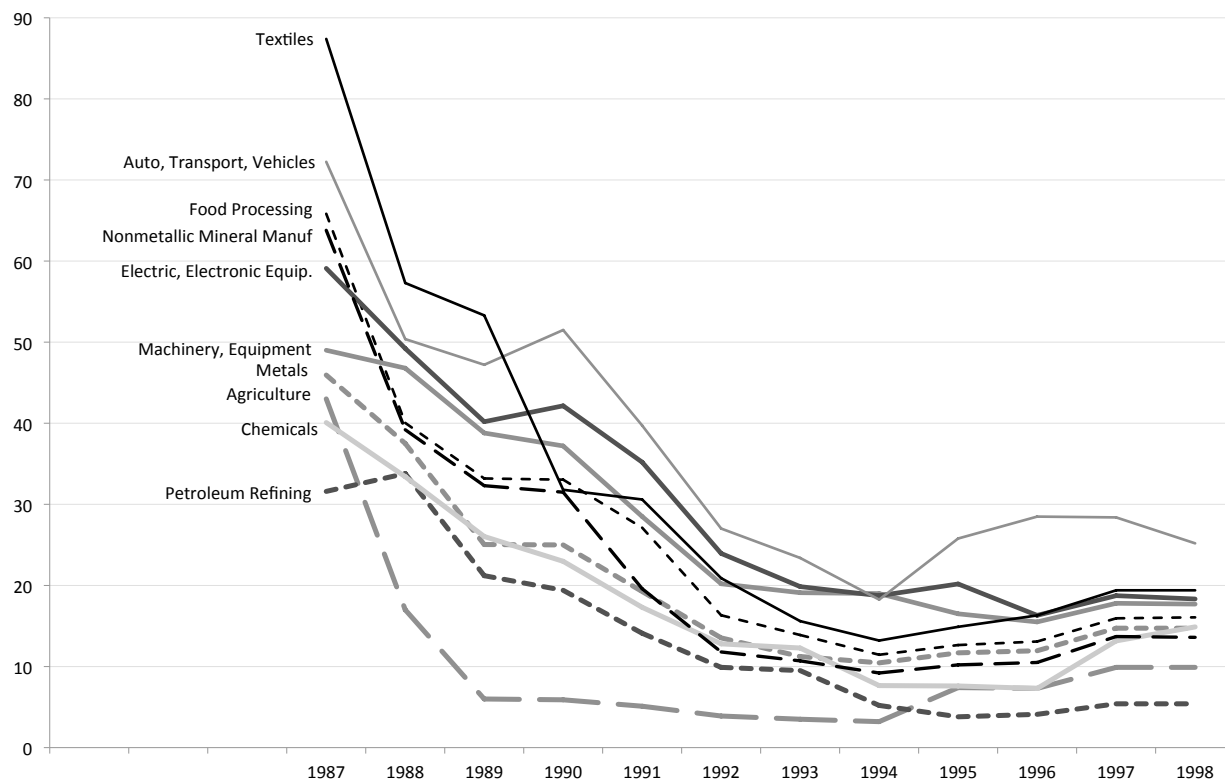
Nominal tariffs from Kume et al. (2003), aggregated to the industry classification presented in Table A1. The ten largest industries by 1990 value added are shown.

A.2   RAIS Data

The *Relação Anual de Informações Sociais* (RAIS) is a high quality census of the Brazilian formal labor market. Originally, RAIS was created as an operational tool for the Brazilian government to i) monitor the entry of foreign workers into the labor market; ii) oversee the records of the FGTS (*Fundo de Garantia do Tempo de Serviço*) program, a national benefits program consisting of employers' contributions to each of its employees; iii) provide information for administering several government benefits programs such as unemployment insurance; and iv) generate statistics regarding the formal labor market. Today it is the main tool used by the government to enable the payment of the "*abono salarial*" to eligible workers. This is a government program that pays one additional minimum wage at the end of the year to workers whose average monthly wage was not greater than two times the minimum wage, and whose job information was correctly declared in RAIS, among other minor requirements. Thus, workers have an incentive to ensure that their employer is filing the required information. Moreover, firms are required to file, and face fines until they do so. Together, these requirements ensure that the data in RAIS are accurate and complete.

Observations in the data are indexed by a worker ID number, the *Programa de Integração Social* (PIS), and an establishment registration number, the *Cadastro Nacional da Pessoa Jurídica* (CNPJ). Both of these identifiers are consistent over time, allowing one to track workers and establishments across years. Establishment industry is reported using the *Subsetor* IBGE classification, which includes 12 manufacturing industries, 2 primary industries, 11 nontradable industries, and 1 other/ignored.[60] Worker education is reported using the following 9 education categories (listing corresponding years of education in parentheses): illiterate (0), primary school dropout (1-3), primary school graduate (4), middle school dropout (5-7), middle school graduate (8), high school dropout (9-10), high school graduate (11), college dropout (12-14), and college graduate (≥ 15).

In each year, and for each job, RAIS reports average earnings throughout the year, and earnings in December.[61] We focus on labor market outcomes reported in December of each year. This choice ensures that earnings and formal employment status are measured at the same time for all workers and all jobs. It avoids the potential confounding effects on average yearly earnings that might arise in situations where some workers begin working in early in the year and others begin late in the year.

A.3   Demographic Census

We utilize information from the long form of the Demographic Censuses (*Censo Demográfico*) for 1970, 1980, 1991, 2000, and 2010. The long form micro data reflect a 5 percent sample of the population in 1970, 1980, and 2010, a 5.8 percent sample in 1991, and a 6 percent sample in 2000. The primary benefit of the Census for our purposes is the ability to observe those outside formal employment, who are not present in the RAIS database.

Although our main analysis focuses on monthly earnings, following the information available in RAIS, the Census provides weekly hours information from 1991-2010, allowing us to calculate hourly wages as monthly earnings divided by 4.33 times weekly hours. Census results for monthly earnings and hourly wages are very similar. In 1970 and 1980, hours information is presented in

---

[60]A less aggregate industry classification (CNAE) is available from 1994 onward, but we need a consistent classification from 1986-2010, so we use *Subsetor* IBGE.

[61]From 1994 onward, RAIS reports hours, making it possible to calculate hourly wages. However, since we need a consistent measure from 1986-2010, we focus on monthly earnings.

5 rough bins. Thus, when calculating pre-liberalization trends using data from 1970 and 1980, we use monthly earnings even when examining hourly wage outcomes.

In 1991-2010, the Census asks whether each worker has a signed work card. This is the standard definition of formal employment, and is necessary for a worker to appear in the RAIS sample. Thus, we use this as our primary definition of formal employment. In 1980 and 1991, there is an alternative proxy for formal employment, reporting whether the worker's job includes contributions to the national social security system. When calculating pre-liberalization outcome trends for 1980-1991, we use this alternative measure to identify formally employed workers. The social security contributions proxy appears to be a good one; in 1991, when both measures are available, 95.9 percent of workers would be classified identically when using either measure. In 1970, there is no information on formality, so pre-liberalization outcome trends for 1970-1980 are calculated for all workers.

The definition of employment changes across Census years. In 1970 it includes those reporting working or looking for work during August 1970 (the questionnaire does not separately identify working vs. looking for work). In 1980 it includes those who report working during the year prior to September 1, 1980. In 1991 it includes those reporting working regularly or occasionally during the year prior to September 1, 1991. In 2000 and 2010 it includes those who report paid work, temporary leave, unpaid work, or cultivation for own consumption during the week of July 23-29 in 2000 and July 25-31 in 2010. Note that the employment concept changes substantially across years. This highlights yet another benefit of using RAIS as our primary data source, since the employment concept in RAIS is consistent throughout the sample. Yet, while the changes complicate the interpretation of Census-based employment rates over time, there is no reason to expect systematic differences across regions to result from the changing employment concept. Thus, our cross-region identification strategy should be valid when using the Census to measure employment in spite of these measurement issues.
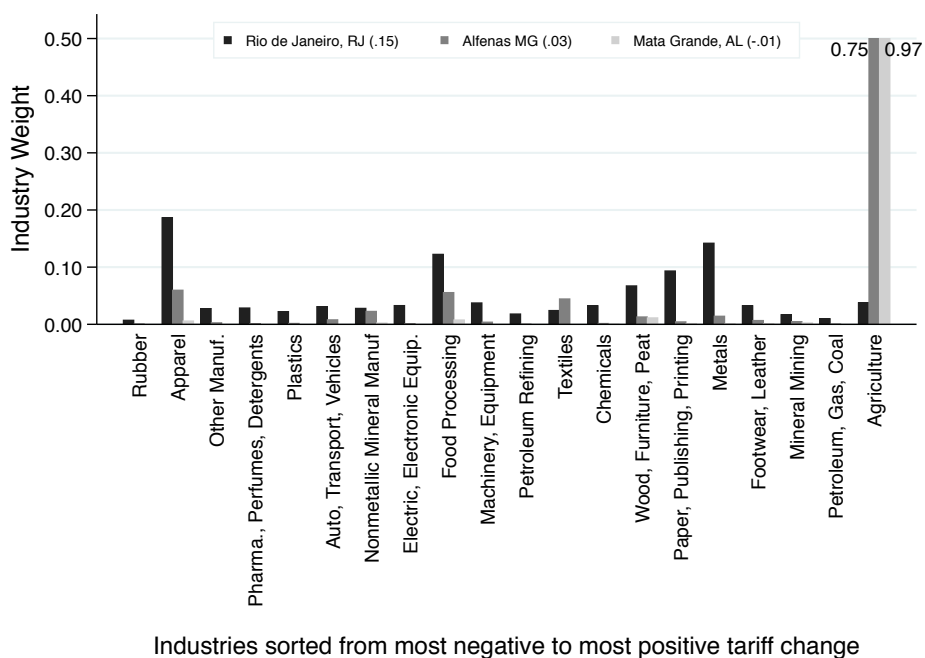
A.4   Regional Tariff Changes

Regional tariff reductions, defined in (2), are constructed using information from various sources. Tariff changes come from Kume et al. (2003), and are aggregated from the *Nível* 50 level to the industry classification presented in Table A1 using 1990 value-added weights from the IBGE National Accounts. Figure 1 shows the resulting industry-level variation in tariff changes.

The weights, β_{ri} in (2) depend upon the initial regional industry distribution (λ_{ri}) and the specific-factor share in production (φ_i). We calculate the λ_{ri} using the 1991 Census. We use the Census because it provides a less aggregate industry definition than what is available in RAIS, and because the Census allows us to calculate weights that are representative of overall employment, rather than just formal employment. However, note that shocks using formal employment weights yield very similar results (see Panel D of Table B6). We calculate the φ_i using data from the Use Table of the 1990 National Accounts from IBGE. The table "*Componentes do Valor Adicionado*" provides the wagebill (*Remunerações*) and gross operating surplus (*Excedente Operacional Bruto Inclusive Rendimento de Autônomos*), which reflects the share of income earned by capital. We define φ_i as capital's share of the sum of these two components. When imposing equal cost shares across industries (see Section 5.4.2), we calculate φ using the economy-wide wagebill and gross operating surplus, yielding a value of $\varphi = 0.544$.

Because Brazilian local labor markets differ substantially in the industry distribution of their

employment, the weights β_{ri} vary across regions. Figure A2 demonstrates how variation in industry mix leads to variation in RTR_r. The figure shows the initial industry distribution of employment for the regions facing the largest tariff reduction (Rio de Janeiro) the median tariff reduction (Alfenas in southwestern Minas Gerais state), and the smallest tariff reduction (actually a small increase, Mata Grande in northwest Alagoas state). The industries on the x-axis are sorted from the most negative to the most positive tariff change. Rio de Janeiro has more weight on the left side of the diagram, by virtue of specializing in manufacturing, particularly in apparel and food processing industries, which faced quite large tariff reductions. Thus, its regional tariff reduction is quite large. Alfenas is a coffee growing and processing region, which also has some apparel employment, balancing the large tariff declines in apparel and food processing against the small tariff increase in agriculture. Mata Grande is located in a sparsely populated mountainous region, and is almost exclusively agricultural, leading it to experience a small tariff increase overall. Thus, although all regions faced the same set of tariff reductions across industries, variation in the industry distribution of employment in each region generates substantial variation in RTR_r.

Figure A2: Variation Underlying Regional Tariff Reduction

Industry distribution of 1991 employment in the regions facing the largest (Rio de Janeiro, RJ), median (Alfenas, MG) and smallest (Mata Grande, AL) regional tariff reduction. Industries sorted from the most negative to the most positive tariff change (see Figure 1). More weight on the left side of the figure leads to a larger regional tariff reduction, and more weight on the right side leads to a smaller regional tariff reduction.

A.5   Local Price Indexes

Moretti (2013) calculates local price indexes for the U.S. using the change in monthly rents for 2 or 3 bedroom apartments. We adjust this approach to the Brazilian context in a few ways. First, we focus on 1 or 2 bedroom apartments, which are far more common in the Brazilian setting, accounting for more than 85 percent of the stock of rental units in 1991 and 2010. Many Brazilian cities include favelas with somewhat improvised structures, and rural areas often feature less formal dwellings. We restrict the sample to include only units with modern construction materials (masonry or wood framing), with at least one bathroom, and with modern sanitation (sewer or septic tank). These restrictions allow us to avoid comparing modern apartments to informal dwellings. Using this sample of apartments, we calculate the change in log average monthly rent in each region. 19 very sparsely populated microregions do not have observations for any rental units satisfying these characteristics in either 1991 or 2010, so we have rent indexes for 456 microregions in our sample.

We then need to transform the change in rental prices into a regional price index. Given the cross-sectional nature of our analysis, we only need to be concerned with prices that vary at the local level, i.e. nontradables, since tradable goods prices move together across regions, and thus do not affect this exercise. Using local Consumer Price Indexes produced by the Bureau of Labor Statistics for 23 U.S. metropolitan areas, Moretti (2013) shows that, as expected, local non-housing nontradables' prices move with local rental prices. He estimates a slope of 0.35 for the effect of housing prices on non-housing nontradables' prices. The Brazilian Consumer Price Index (*Índices de Preços ao Consumidor* - IPC) system reports that in 2002-03, housing's share of consumption was 16.24 percent and that the share for other nontradable goods was 39.94 percent (IBGE 2005). Together, these figures imply that the effective weight on housing prices in the consumer price index is $0.1624 + 0.3994 \cdot 0.35 = 0.3022$. Our local price deflator is therefore 0.3022 times the change in log rental prices in the region.

A.6 Regional Change in Imports and Exports

Import and export data between Brazil and the rest of the world come from Comtrade, at the 6-digit HS level. We map from HS codes to the industries presented in Table A1 and calculate total Brazilian trade flows by industry and year.

In the main text, we follow Autor et al. (2013) (ADH) by generating regional weighted averages of changes in imports and exports per worker. For each industry, we calculate the change in imports (M_{it}) and exports (X_{it}) from 1990 to each subsequent year t. These trade flows are measured in $100,000 units. We then generate the regional change in imports and exports per worker as follows.

$$RegImp_{rt} = \sum_i \frac{L_{rit_0}}{L_{rt_0}} \frac{\Delta_{1990-t} M_{it}}{L_{it_0}} \tag{16}$$

$$RegExp_{rt} = \sum_i \frac{L_{rit_0}}{L_{rt_0}} \frac{\Delta_{1990-t} X_{it}}{L_{it_0}} \tag{17}$$

The rightmost ratios in these expressions measure the change in imports or exports per worker initially employed in the industry, in year $t_0 = 1991$. The preceding ratios represent industry weights for each region, reflecting industry i's share of tradable employment in region r in 1991. These weights are equivalent to λ_{ri} in (1). We then generate weighted averages by summing these terms over tradable industries. Finally, we construct regional net exports as the difference in regional exports and imports.

$$RegNetExp_{rt} = RegExp_{rt} - RegImp_{rt}$$

In Appendix B.10, we present an alternative set of results based on the change in log trade flows rather than the change in trade flows per worker.

$$RegLnImp_{rt} = \sum_i \frac{L_{rit_0}}{L_{rt_0}} \Delta_{1990-t} \ln(M_{it}) \tag{18}$$

$$RegLnEmp_{rt} = \sum_i \frac{L_{rit_0}}{L_{rt_0}} \Delta_{1990-t} \ln(X_{it}) \tag{19}$$

We emphasize that this measure is presented only for descriptive purposes and as a statistical robustness test since it does not have the same theoretical underpinnings as the measures following Autor et al. (2013).

A.7 Entry, Exit, Job Creation, and Job Destruction

We calculate cumulative job creation and job destruction following Davis and Haltiwanger (1990).

$$job\_creation_{rt} \equiv \sum_{e \in E_{rt}, \; g_{et} > 0} \frac{x_{et}}{X_{rt}} g_{et}, \tag{20}$$

$$job\_destruction_{rt} \equiv \sum_{e \in E_{rt}, \; g_{et} < 0} \frac{x_{et}}{X_{rt}} |g_{et}|, \tag{21}$$

10

$$\text{where} \quad g_{et} \equiv \frac{L_{et} - L_{e,1991}}{x_{et}}, \quad x_{et} \equiv \frac{1}{2}(L_{et} + L_{e,1991}), \quad X_{rt} \equiv \sum_{e \in E_{rt}} L_{et},$$

L_{et} is employment at establishment e in year t and E_{rt} is the set of active establishments in region r in year t. Note that employment growth, g_{et}, is calculated from 1991 to year t. The dependent variables for the regressions underlying Figure 9 are $\ln(job\_creation_{rt})$ and $\ln(job\_destruction_{rt})$.

Entry and exit are calculated analogously, replacing establishment employment with an indicator for the establishment being active in the relevant year, ι_{et}.

$$entry_{rt} \equiv \sum_{e \in E_{rt}, \ \tilde{g}_{et} > 0} \frac{\tilde{x}_{et}}{\tilde{X}_{rt}} \tilde{g}_{et}, \tag{22}$$

$$exit_{rt} \equiv \sum_{e \in E_{rt}, \ \tilde{g}_{et} < 0} \frac{\tilde{x}_{et}}{\tilde{X}_{rt}} |\tilde{g}_{et}|, \tag{23}$$

$$\text{where} \quad \tilde{g}_{et} \equiv \frac{\iota_{et} - \iota_{e,1991}}{\tilde{x}_{et}}, \quad \tilde{x}_{et} \equiv \frac{1}{2}(\iota_{et} + \iota_{e,1991}), \quad \tilde{X}_{rt} \equiv \sum_{e \in E_{rt}} \iota_{et},$$

These definitions yield intuitive results. $entry_{rt}$ is equivalent to the share of active firms in region r in year t that were not active in 1991. $exit_{rt}$ is equivalent to the number of firms in region r that were active in 1991 and not in year t, divided by the number of firms active in year t. The dependent variables for the regressions underlying Figure 8 are $\ln(entry_{rt})$ and $\ln(exit_{rt})$

B   Supplemental Empirical Results

B.1   Industry-Level Outcome Pre-Trends vs. Tariff Reductions

Along with regional variation in the industrial composition of employment, our analysis relies on variation in tariff cuts across industries. Here we analyze the relationship between tariff cuts during liberalization (1990-1995) and trends in industry wages and employment before liberalization, 1980-1991. We calculate these pre-liberalization outcome trends using the Demographic Census, to provide a longer pre-liberalization period than what is available in RAIS, which starts in 1986.

We implemented a variety of specifications, with results reported in Table B1. In all specifications, the independent variable is the proportional reduction in one plus the tariff rate ($-\Delta_{1990-95} \ln(1 + \tau_i)$). In panels A-C the dependent variable is the change in log industry earnings. Panel A uses average log earnings; Panel B uses average log earnings residuals controlling for individual age, sex, education, and formal status; and Panel C uses average log earnings residuals controlling for these individual characteristics and region fixed effects. In Panel D, the dependent variable is the change in industry log employment. Column (1) weights industries equally, and presents standard errors based on pairwise bootstrap of the t-statistic, to improve small sample properties with only 20 tradable industry observations. Column (2) uses the same estimator, but drops agriculture. Column (3) uses heteroskedasticity weights and presents heteroskedasticity-robust standard errors, which are likely understated in this small sample (MacKinnon 2011). Column (4) uses the same estimator, but drops agriculture. In all cases, the results should be seen primarily as suggestive, because the analysis uses only 19 or 20 observations.

Nearly all of the earnings estimates are positive, indicating larger tariff reductions in industries experiencing more positive wage growth prior to liberalization. The majority of the estimates are insignificantly different from zero, with the exception of weighted results in Panels A and B. These specifications heavily weight agriculture, which exhibited negative wage growth prior to liberalization and experienced essentially no tariff decline during liberalization, driving the strong negative relationship. By dropping agriculture, Column (4) confirms that the significant relationship is driven by agriculture. The employment estimates are larger, and change sign across columns. Given the diversity of findings across earnings and employment specifications, this exercise is somewhat inconclusive. Tariff cuts may or may not have been substantially correlated with pre-liberalization outcome trends. These findings motivate us to control for pre-liberalization outcome trends whenever possible throughout the paper. This ensures that our results are robust to potential spurious correlation between liberalization-induced labor demand shocks and ongoing trends.

Table B1: Pre-Liberalization Industry Trends - 1980-1991

1980-1991 change in log:	unweighted, bootstrapped (1)	unweighted, bootstrapped, omitting agriculture (2)	weighted (3)	weighted, omitting agriculture (4)
Panel A: average earnings				
Industry tariff reduction	0.345	0.111	1.029***	0.510
	(0.322)	(0.354)	(0.139)	(0.582)
Panel B: earnings premia (with individual controls)				
Industry tariff reduction	0.203	-0.017	0.610***	-0.235
	(0.273)	(0.311)	(0.157)	(0.350)
Panel C: earnings premia (with individual and region controls)				
Industry tariff reduction	0.135	0.044	0.184	0.018
	(0.177)	(0.209)	(0.158)	(0.222)
Panel D: employment				
Industry tariff reduction	-1.624	-2.696**	0.687	-1.651
	(1.272)	(1.361)	(0.417)	(1.894)
Observations	20	19	20	19

Decennial Census data. 20 industry observations (19 omitting agriculture). See text for details of dependent and independent variable construction. Column (1) weights industries equally, and presents standard errors based on pairwise bootstrap of the t-statistic. Column (2) uses the same estimator as Column (1), but drops agriculture. Column (3) uses heteroskedasticity weights and presents heteroskedasticity-robust standard errors. Column (4) uses the same estimator as Column (3), but drops agriculture. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.2   Informal Employment

The following results provide some descriptive evidence on the informal sector in Brazil. Informality is defined as working without a signed work card (*Carteira de Trabalho e Previdência Social*), which entitles workers to benefits and labor protections afforded them by the legal employment system. Table B2 shows that the overall rate of informality increased from 1991 to 2000, before decreasing substantially from 2000 to 2010. Rates of informality are highest in agriculture and much lower in manufacturing. Table B1 breaks out informality rates in the manufacturing sector into individual industries. Finally, Table B2 focuses on the year 2000 and shows the industry distribution of formal and informal employment. There is very substantial overlap in the industry distributions of formal and informal employment. The biggest differences occur in agriculture, which comprises a much larger share of informal employment, and food processing and metals, which comprise larger shares of formal employment. In contrast, the nontradable share is nearly identical for formal and informal employment.

Table B2: Informal Share of Employment - 1991-2010

	1991	2000	2010
Overall	0.58	0.64	0.49
Agriculture	0.89	0.86	0.83
Mining	0.61	0.45	0.21
Manufacturing	0.28	0.39	0.29
Nontradable	0.55	0.64	0.48

Author's calculations using Brazilian Demographic Census data for workers age 18-64. Informality defined as not having a signed work card.

Figure B1: Informal Share of Employment by Industry - 1991-2010
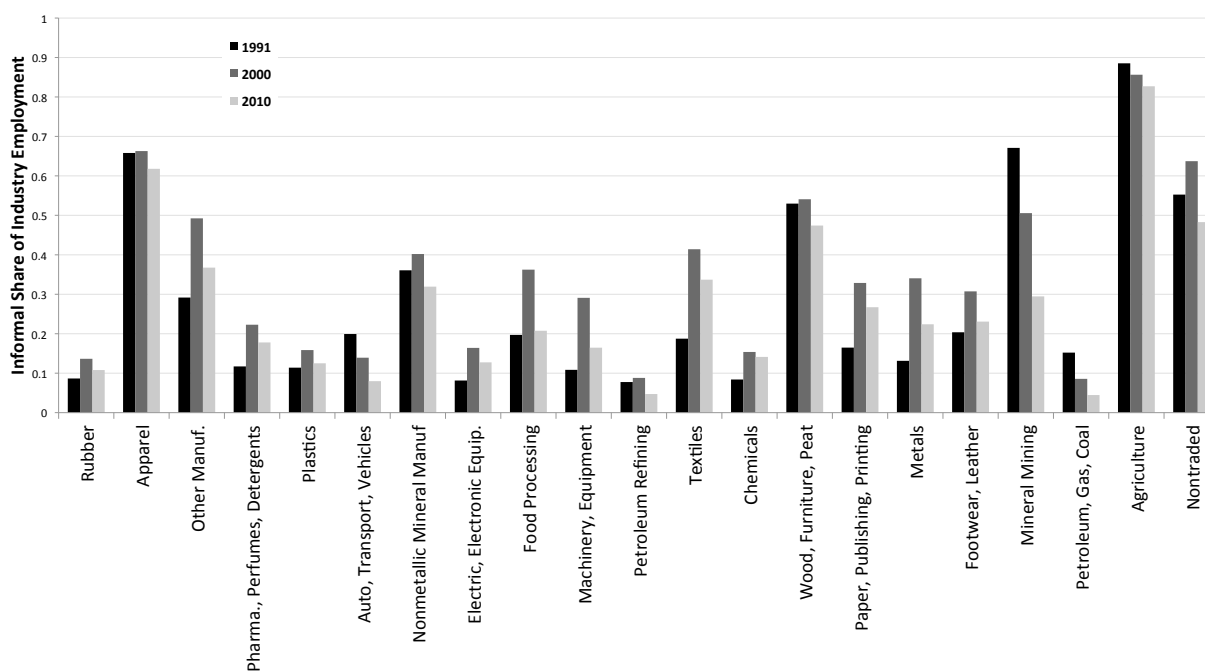
Authors' calculations using Brazilian Demographic Census data for workers age 18-64. Informality defined as not having a signed work card. Industries sorted from most negative to most positive tariff change (with the exception of the nontraded sector).

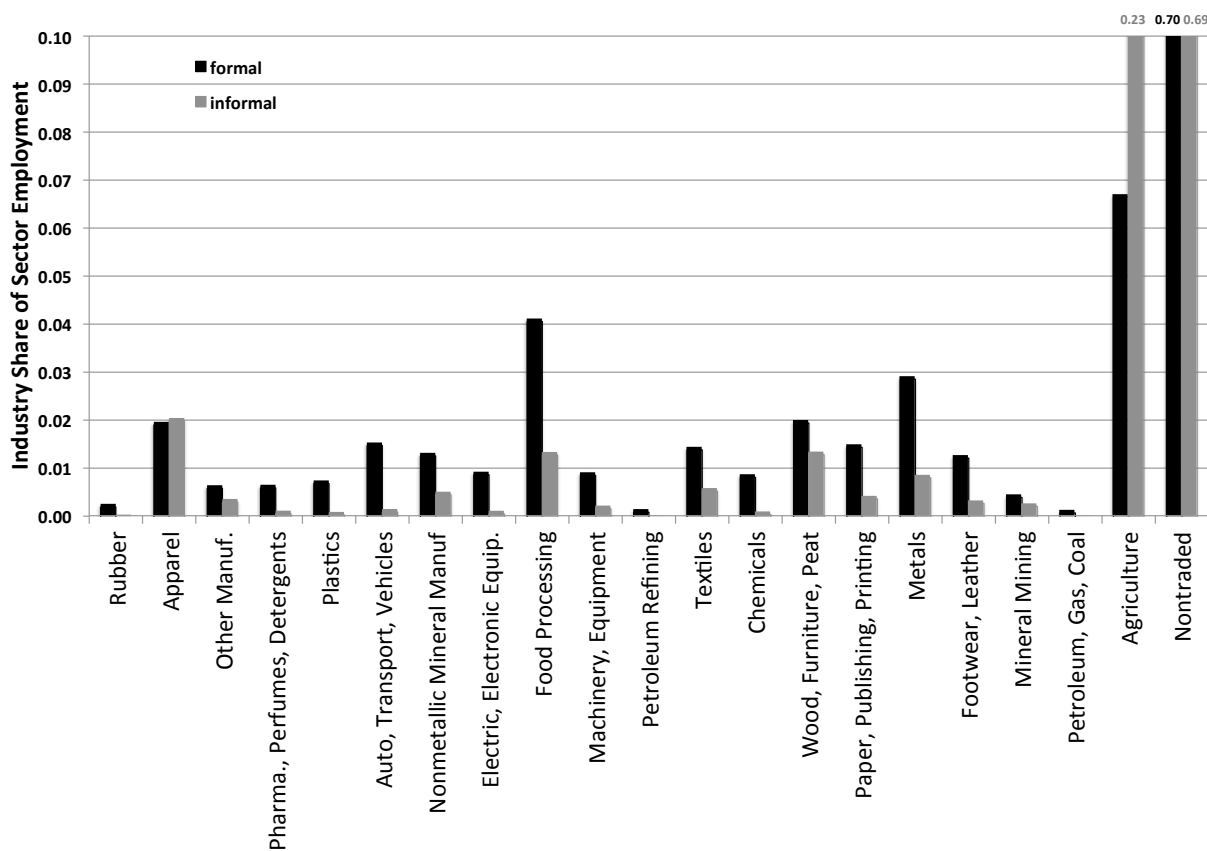Figure B2: Industry Distribution of Formal and Informal Employment - 2000

Authors' calculations using year 2000 Brazilian Demographic Census data for workers age 18-64. Informality defined as not having a signed work card. Industries sorted from most negative to most positive tariff change (with the exception of the nontraded sector).

B.3   Regional Earnings Premium Regressions

As discussed in Section 4.1, we calculate regional earnings premia by regressing workers' log December earnings on flexible demographic and educational controls, industry fixed effects, and region fixed effects, separately in each year. Table B3 shows the coefficient estimates from these earnings premium regressions for 1991, 2000, and 2010. The region fixed effect estimates provide average log earnings for formally employed workers in the region, controlling for the age, sex, education, and industry composition of the region's employment. These regional premia then form the outcome variable in our earnings analyses.

Note that the coefficient estimates on the controls conform to expectations. Women are paid less than otherwise similar men, and this earnings gap declines over time. Workers exhibit an inverted U-shaped wage profile as they age, as is standard in Mincerian regressions. The returns to education are monotonically positive.

Table B3: Regional Earnings Premium Regressions - 1991, 2000, 2010

	dependent variable: log monthly earnings		
	(1) 1991	(2) 2000	(3) 2010
Female	-0.349***	-0.295***	-0.261***
	(0.000)	(0.000)	(0.000)
Age			
25-29	0.210***	0.209***	0.148***
	(0.000)	(0.000)	(0.000)
30-39	0.407***	0.374***	0.273***
	(0.000)	(0.000)	(0.000)
40-49	0.527***	0.525***	0.382***
	(0.001)	(0.000)	(0.000)
50-64	0.435***	0.507***	0.474***
	(0.001)	(0.001)	(0.000)
Education (years)			
Primary School Dropout (1-3)	0.015***	0.007***	0.130***
	(0.001)	(0.001)	(0.001)
Primary School Graduate (4)	0.111***	0.075***	0.182***
	(0.001)	(0.001)	(0.001)
Middle School Dropout (5-7)	0.188***	0.129***	0.206***
	(0.001)	(0.001)	(0.001)
Middle School Graduate (8)	0.297***	0.181***	0.236***
	(0.001)	(0.001)	(0.001)
High School Dropout (9-11)	0.454***	0.305***	0.289***
	(0.001)	(0.001)	(0.001)
High School Graduate (12)	0.711***	0.523***	0.430***
	(0.001)	(0.001)	(0.001)
College Dropout (13-15)	0.967***	0.902***	0.792***
	(0.002)	(0.001)	(0.001)
College Graduate ($\geq$16)	1.374***	1.384***	1.368***
	(0.001)	(0.001)	(0.001)
Fixed Effects			
Industry (24)	X	X	X
Region (475)	X	X	X
Observations	13,582,443	17,733,492	30,662,075
R-squared	0.858	0.842	0.759

Individual worker observations from RAIS. Earnings premium regressions were run for each year from 1986-2010. Here we show three years as examples. The region fixed effect estimates provide average log earnings for formally employed workers in the region, controlling for the age, sex, education, and industry composition of the region's employment. These regional premia then form the outcome variable in our regional earnings analyses. The omitted category is a male, age 18-24, with 0 years of education (illiterate). Robust standard errors in parentheses. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.4   Industry-Region Earnings Results

In this appendix, we pursue an alternative research design for studying the earnings effects of liberalization, with the unit of analysis at the industry $\times$ region level. Equations (13) and (14) suggest a similar industry $\times$ region research design, but in the context of our model this approach only applies in the long run. The results in this section are therefore supplementary. Similar approaches appear in Hakobyan and McLaren (forthcoming) and Acemoglu, Autor, Dorn, Hanson and Price (2016).

We must first estimate earnings premia at the industry $\times$ region level. Because there are 4,773 such combinations, it is not feasible to directly estimate industry $\times$ region fixed effects (and obtain their standard errors) in an earnings regression like (4). Instead, we estimate the following earnings regression separately in each year t, absorbing the industry $\times$ region fixed effects, μ_{irt}

$$\ln(earn_{jrit}) = X_{jt}\Gamma_t + \mu_{irt} + e_{jrit} \tag{24}$$

We then calculate residuals $\hat{u}_{jrit} \equiv \ln(earn_{jrit}) - X_{jt}\hat{\Gamma}_t$, and average these residuals within industry $\times$ region bins to recover the fixed-effect estimates, $\hat{\mu}_{irt}$. Because this procedure does not directly yield standard errors for these fixed-effect estimates, we implement 100 bootstrap repetitions and calculate bootstrap standard errors. We then calculate the change in these industry $\times$ region earnings premia and use them as the dependent variable in the following regression specification.

$$\hat{\mu}_{irt} - \hat{\mu}_{ir,1991} = \theta_t RTR_r + \delta_t d\ln(1+\tau_i) + \alpha_{st} + \gamma_t(\hat{\mu}_{ir,1990} - \hat{\mu}_{ir,1986}) + \varepsilon_{irt} \tag{25}$$

We estimate this regression separately in each year $t > 1991$. Note that the unit of observation is the industry $\times$ region, and we include both the region-level tariff reduction, RTR_r, and the industry-level tariff change, $d\ln(1+\tau_i)$. In the most stringent specifications, we control for state fixed effects and both region and industry earnings pre-trends.

The results for 2000 and 2010 appear in Panel A of Table B4. First, note that the estimated regional and industry-level effects of liberalization are quite consistent across specifications. We continue to find large increases in the regional effects of liberalization between 2000 and 2010, while the industry-level results are quite constant over time. The industry-level tariff change estimates reflect the direct effect of liberalization on workers in the affected industry, irrespective of their region of residence. The RTR_r coefficients capture the regional general-equilibrium effects operating across industries. In 2010, the coefficient on RTR_r is more than three times larger than the industry effects. This finding makes clear the central role of regional labor market equilibrium in affecting workers outcomes following trade liberalization. Panel B replaces the controls for the industry-level tariff change with industry fixed effects, finding similar results.

Table B4: Industry × Region log Formal Earnings 2000, 2010

Change in industry × region earnings	1991-2000			1991-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Industry Tariff-Change Controls						
Regional tariff reduction (RTR)	-0.618**	-0.878***	-0.762***	-2.567***	-2.485***	-2.252***
	(0.270)	(0.276)	(0.254)	(0.582)	(0.266)	(0.216)
Industry tariff reduction (-$d \ln(1+\tau_i)$)	-0.683***	-0.654***	-0.661***	-0.684***	-0.573**	-0.590***
	(0.215)	(0.216)	(0.215)	(0.254)	(0.251)	(0.215)
Formal earnings pre-trend (86-90)			-0.196**			-0.393***
			(0.080)			(0.075)
State fixed effects (26)		✓	✓		✓	✓
R-squared	0.082	0.116	0.143	0.212	0.323	0.390
Panel B: Industry Fixed Effects						
Regional tariff reduction (RTR)	-0.528**	-0.777***	-0.724***	-2.199***	-2.109***	-2.036***
	(0.229)	(0.217)	(0.212)	(0.538)	(0.248)	(0.244)
Formal earnings pre-trend (86-90)			-0.243***			-0.277***
			(0.034)			(0.035)
State fixed effects (26)		✓	✓		✓	✓
R-squared	0.313	0.345	0.375	0.437	0.530	0.554

Negative coefficient estimates for the regional tariff reduction imply larger declines in formal earnings in regions facing larger tariff reductions. Negative coefficients for the industry tariff reduction imply larger declines in formal earnings in industries facing larger tariff reductions. 4,733 industry × region observations in each year. Earnings premia calculated controlling for age, sex, and education. Panel A controls for industry tariff reductions, while Panel B uses industry fixed effects to capture liberalization's direct effect on each industry. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium (see text for description of bootstrap standard error calculation). Pre-trends computed for 1986-1990. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.5   Formal Earnings Regression Scatterplots

Figure B3 shows scatter plots underlying the formal earnings regression estimates in Figure 3 for 1995, 2000, 2005, and 2010. Each marker represents a microregion, and microregions in each major region are shown with a separate type of marker. The size of each marker is proportional to the weight the relevant microregion receives in the estimation. The mean value of the dependent variable is normalized to zero in each year to focus attention on the slope.

These scatter plots make clear three important points about the earnings estimates. First, as shown in Figure 3, the magnitude of the slope increases substantially and steadily as time passes following liberalization. Second, the relationship between changes in formal earnings premia and regional tariff reductions is approximately linear in all time periods, justifying our choice of functional form. Third, the increasing magnitude slope is driven by shifts in earnings across large numbers of microregions in various parts of the country, rather than by a few outliers.

Figure B3: Formal Earnings Premia vs. Regional Tariff Reductions – 1995, 2000, 2005, 2010
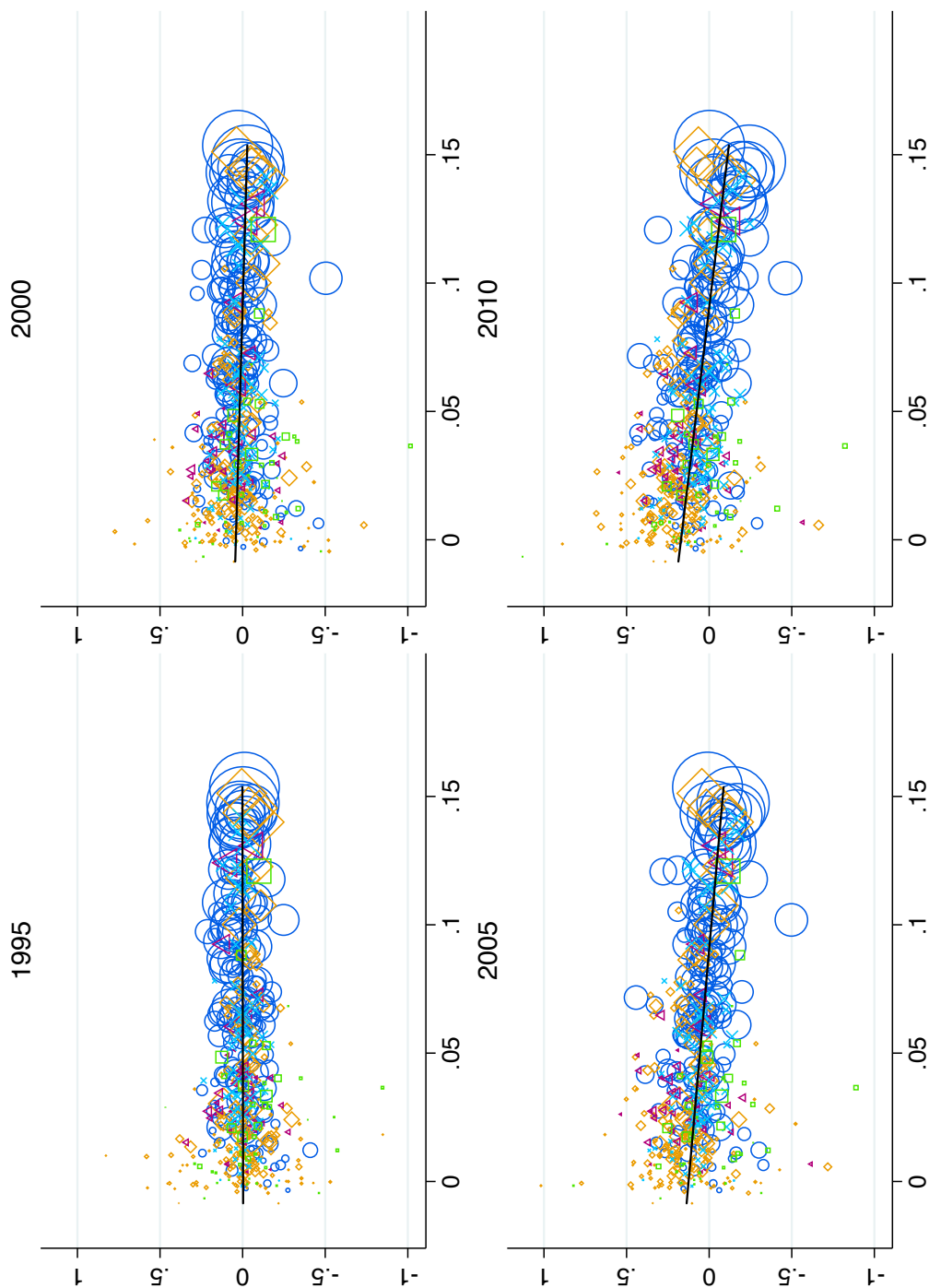
Scatter plots of regressions behind Figure 3 estimates. 475 microregion observations. Y-axis is the change in log formal earnings premium in a given region from 1991 to the year listed in the title. Dependent variable demeaned in each year. X-axis is the regional tariff reduction, RTR_r. The size of each marker represents the inverse of the variance of the dependent variable estimate, used to weight the regressions. Major regions shown with different symbols. Southeast: dark blue circles, Northeast: orange diamonds, Center-West: red triangles, South: light blue Xs, North: green squares.

B.6   Census Earnings, Wage, and Employment Results

Table B5 estimates versions of equation (3) using formal sector outcomes calculated using Census data. Because the Census includes hours information, we are able to examine both earnings and wage premia. In all cases, we find negative coefficients on RTR_r, indicating that regions facing larger tariff reductions experienced relative declines in monthly earnings, hourly wages, or employment. We also find substantial growth in the magnitude of these effects, corroborating the results in Table 2, which uses RAIS outcomes.

Table B5: Census Regional log Formal Earnings, Wages, and Employment - 2000, 2010

Change in outcome:		1991-2000			1991-2010	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: log Formal Earnings Premia						
Regional tariff reduction (RTR)	-0.397	-0.293**	-0.261**	-1.384**	-0.890***	-0.855***
	(0.335)	(0.120)	(0.116)	(0.572)	(0.198)	(0.186)
Formal earnings pre-trend (86-90)			-0.0896*			-0.0994
			(0.0528)			(0.0728)
State fixed effects (26)		✓	✓		✓	✓
R-squared	0.031	0.579	0.583	0.156	0.718	0.720
Panel B: log Formal Wage Premia						
Regional tariff reduction (RTR)	-0.630*	-0.533***	-0.495***	-1.320**	-0.765***	-0.721***
	(0.355)	(0.124)	(0.118)	(0.525)	(0.173)	(0.163)
Formal earnings pre-trend (86-90)			-0.108*			-0.124*
			(0.0555)			(0.0642)
State fixed effects (26)		✓	✓		✓	✓
R-squared	0.071	0.605	0.609	0.136	0.718	0.721
Panel B: log Formal Employment						
Regional tariff reduction (RTR)	-2.478***	-1.756***	-1.619***	-3.913***	-2.865***	-2.725***
	(0.487)	(0.281)	(0.258)	(0.758)	(0.443)	(0.417)
Formal employment pre-trend (86-90)			0.211***			0.227**
			(0.0704)			(0.0926)
State fixed effects (26)		✓	✓		✓	✓
R-squared	0.317	0.612	0.630	0.319	0.629	0.638

Outcomes calculated using Census data. Negative coefficient estimates for the regional tariff reduction imply larger declines in formal earnings, wages, or employment in regions facing larger tariff reductions. 475 microregion observations. Regional earnings premia calculated controlling for age, sex, education, and industry of employment. Efficiency weighted by the inverse of the squared standard error of the estimated outcome. RAIS pre-trends computed for 1986-1990. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.7   Robustness Tests

Table B6 presents the various earnings robustness tests summarized in Section 4.2. For comparison, Panel A shows the main specification, corresponding to the estimates in Table 2 and Figure 3.

All panels in the table control for pre-liberalization earnings growth from 1986-1990, calculated using RAIS. Panel B additionally controls for longer pre-liberalization earnings trends calculated using the Census. The 1980-1991 control reflects the growth in formal earnings premium, where formality is defined based on whether the worker's job included social security contributions. The 1970-1980 control reflects growth in the earnings premium for all workers, since there is no formality information in the 1970 Census. See Appendix A.3 for more detail on Census data. A potential problem with Panel B is mechanical endogeneity, because the 1980-1991 pre-trend and the 1991-t earnings growth dependent variable both include the 1991 earnings premium. Panel C resolves this issue by using earnings growth from 1992 to year t as the dependent variable, while including the Census pre-trend controls.

Panel D calculates the regional tariff reduction (RTR_r) in (2) using weights based on the initial industry distribution of regional *formal* employment, rather than overall employment. Panel E calculates RTR_r using effective rates of protection rather than nominal tariffs. Effective rates of protection capture the overall effect of liberalization on producers in a given industry, accounting for tariff changes on industry inputs and outputs. Kume et al. (2003) provide effective rates of protection along with the nominal tariffs used in our main analysis. The magnitude of the changes in effective rates of protection is larger than for nominal tariffs, so the coefficients in Panel 3 smaller by the same proportion. Since versions of RTR_r based on effective rates of protection and nominal tariffs are nearly perfectly correlated (correlation = 0.993), the variation in earnings growth explained by both versions is nearly identical. Panel F calculates RTR_r including the nontradable sector with a tariff change value of 0. This measure ignores the fact that nontradable prices move with tradable prices (see Appendix A.5 and Panel B of Table 2), and in doing so underestimates the magnitude of the average liberalization-induced price change faced by each region. Because the magnitude of RTR_r is reduced, the coefficient estimates are inflated by the same proportion.

Panel G omits industry fixed effects when calculating regional earnings premia. This maintains the national industry-level variation in earnings in the outcome measure, rather than restricting attention to the regional equilibrium earnings used in our main specifications. Panel H omits all controls from the earnings premium regressions, using simple average log earnings for workers in the relevant region. While the main analysis weights by the inverse of the squared standard error of the estimated growth in regional wage premium, Panel I weights all regions equally, and Panel J weights by 1991 formal employment.

In all cases, the effects grow substantially over time, as in our main specification. In fact, in all but two of these robustness tests, the long run effect of liberalization on earnings is *larger* than it is in the main specification.

Table B7 shows that the formal employment results in Table 2 and Figure 4 are similarly robust. The panel labels correspond to Table B6, so see above for descriptions of each specification. Note that Panels G and H do not apply to employment, since they relate to earnings premia. Panel I also does not apply, because the main employment specification is unweighted, since RAIS contains the population of formally employed workers.

Table B6: Robustness: Regional log Formal Earnings Premia - 1995, 2000, 2005, 2010

Change in log Formal Earnings Premia:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Main specification				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Panel B: Long pre-trends (Census: 1970-80, 1980-91, RAIS: 1986-90)				
Regional tariff reduction (RTR)	-0.243*	-0.770***	-1.498***	-1.814***
	(0.130)	(0.184)	(0.186)	(0.199)
Panel C: Long pre-trends, earnings growth from 1992 to t				
Regional tariff reduction (RTR)	-0.478***	-1.009***	-1.737***	-2.039***
	(0.115)	(0.197)	(0.195)	(0.215)
Panel D: RTR using formal employment industry weights				
Regional tariff reduction (RTR)	-0.089	-0.358	-1.270***	-1.665***
	(0.189)	(0.217)	(0.246)	(0.298)
Panel E: RTR using effective rates of protection				
Regional tariff reduction (RTR)	-0.047	-0.328***	-0.823***	-1.017***
	(0.076)	(0.091)	(0.090)	(0.107)
Panel F: RTR including zero nontradable price change				
Regional tariff reduction (RTR)	-0.798	-1.758***	-3.350***	-4.625***
	(0.489)	(0.577)	(0.643)	(0.696)
Panel G: Earnings premium without industry fixed effects				
Regional tariff reduction (RTR)	0.131	-0.422***	-1.420***	-1.895***
	(0.148)	(0.151)	(0.163)	(0.209)
Panel H: Earnings premium with no controls (mean log earnings)				
Regional tariff reduction (RTR)	0.317*	0.046	-1.192***	-1.905***
	(0.171)	(0.214)	(0.147)	(0.182)
Panel I: Unweighted (equally weighted)				
Regional tariff reduction (RTR)	-0.244	-0.490**	-1.074***	-1.546***
	(0.169)	(0.197)	(0.215)	(0.224)
Panel J: Weighted by 1991 formal employment				
Regional tariff reduction (RTR)	-0.020	-0.345**	-1.217***	-1.631***
	(0.128)	(0.147)	(0.143)	(0.175)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. 475 microregion observations, except Panels B and C, which use a more aggregate region definition with 405 observations for consistency with 1970 and 1980 Census data. Regional earnings premia calculated controlling for age, sex, education, and industry of employment except in Panels G and H. Standard errors (in parentheses) adjusted for 112 mesoregion clusters, except Panels B and C with 90 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium except in Panels I and J. See text for detailed description of each panel. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

Table B7: Robustness: Regional log Formal Employment - 1995, 2000, 2005, 2010

Change in log Formal Employment:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Main specification				
Regional tariff reduction (RTR)	-1.900***	-3.533***	-4.517***	-4.663***
	(0.422)	(0.582)	(0.685)	(0.679)
Panel B: Long pre-trends (Census: 1970-80, 1980-91, RAIS: 1986-90)				
Regional tariff reduction (RTR)	-1.157	-3.393***	-4.687***	-4.537***
	(0.787)	(0.930)	(1.019)	(1.007)
Panel C: Long pre-trends, earnings growth from 1992 to t				
Regional tariff reduction (RTR)	-0.722	-2.957***	-4.252***	-4.102***
	(0.804)	(0.972)	(1.084)	(1.070)
Panel D: RTR using formal employment industry weights				
Regional tariff reduction (RTR)	-1.728***	-2.690***	-4.491***	-4.362***
	(0.598)	(0.793)	(0.782)	(0.789)
Panel E: RTR using effective rates of protection				
Regional tariff reduction (RTR)	-1.201***	-2.336***	-2.959***	-3.074***
	(0.274)	(0.369)	(0.438)	(0.430)
Panel F: RTR including zero nontradable price change				
Regional tariff reduction (RTR)	-5.677***	-8.574***	-10.874***	-12.507***
	(1.571)	(2.441)	(2.752)	(2.750)
Panel G: Not applicable				
Panel H: Not applicable				
Panel I: Not applicable				
Panel J: Weighted by 1991 formal employment				
Regional tariff reduction (RTR)	-1.195***	-2.119***	-3.406***	-2.842***
	(0.195)	(0.358)	(0.340)	(0.397)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal employment in regions facing larger tariff reductions. 475 microregion observations, except Panels B and C, which use a more aggregate region definition with 405 observations for consistency with 1970 and 1980 Census data. Panel labels correspond to Table B6, so Panels G and H, which relate to earnings premia, are not applicable here, nor is Panel I, since the main specification is unweighted. Standard errors (in parentheses) adjusted for 112 mesoregion clusters, except Panels B and C with 90 mesoregion clusters. See text for detailed description of each panel. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.8   Potential Confounders

B.8.1   Post-Liberalization Tariff Reductions

We calculate post-liberalization regional tariff reductions as in (2), but use tariff reductions between 1995 and year $t > 1995$. Because the Kume et al. (2003) data end in 1998, we use UNCTAD TRAINS to construct post-liberalization tariff reductions. The TRAINS data are reported by 6-digit HS codes. In order to maintain as much industry variation as possible, we created an industry mapping from HS codes to Census industry codes, which yields 44 consistently identifiable tradable industries. This provides more industry detail than the main industry definition in Table A1. The concordance is available upon request. Panel B of Table B8 includes these post-liberalization tariff reduction controls in the regional earnings growth regression. The post-liberalization control has the expected negative coefficient, but its inclusion has very little effect on the liberalization coefficient.

B.8.2   Real Exchange Rates

We construct regional real exchange rate shocks as follows. We begin with real exchange rates between Brazil and its trading partners, calculated from Revision 7.1 of the Penn World Tables. We then calculate each country's 1989 shares of Brazil's imports and exports in each industry using Comtrade. As in the prior section, we use the industry definition mapping from HS codes to Census industries. Industry-specific real exchange rates are weighted averages of country-specific real exchange rates, weighting either by the 1989 import share or export share. We define industry-level real exchange rate shocks as the change in log industry real exchange rate from 1990 to each subsequent year. Finally we create regional real exchange rate shocks as weighted averages of industry real exchange rate shocks, where the region's industry weights are given by the 1991 industry distribution of employment. Panel C of Table B8 includes both the import-weighted and export-weighted real exchange rate controls. With these controls, the earnings effects grow even more than in the main specification.

B.8.3   Privatization

Substantial privatization in Brazil began in 1991 with the administration of President Collor, but significantly increased during President Cardoso's administration (1995-2002). Beginning in 1995, the RAIS data allow us to identify as state-owned any firm at least partly owned by the government. In panels D and E of Table B8, we include different controls for the regional effects of privatization. Panel D includes quartile indicators for the 1995 share of regional employment in state-owned firms, controlling flexibly for the initial share of employment subject to potential privatization. Panel E controls for the change in state-owned firm employment share from 1995 to subsequent year t. In both cases, the privatization controls have no meaningful effect on the RTR_r coefficients.

Table B8: Potential Confounders: Regional log Formal Earnings Premia - 1995, 2000, 2005, 2010

Change in log Formal Earnings Premia:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Main specification				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Panel B: Post-liberalization tariff reductions				
Regional tariff reduction (RTR)	-0.096	-0.542***	-1.234***	-1.809***
	(0.120)	(0.137)	(0.171)	(0.218)
Post-liberalization (1995 to t) regional	n/a	-3.705	-2.415	-2.124
tariff reductions		(3.273)	(3.872)	(1.425)
Panel C: Exchange rates				
Regional tariff reduction (RTR)	-0.113	-0.482***	-1.475***	-1.728***
	(0.118)	(0.161)	(0.184)	(0.207)
Import-weighted real exchange rate	0.136	0.570*	0.243*	0.374*
	(0.133)	(0.327)	(0.141)	(0.201)
Export-weighted real exchange rate	0.051	-0.164	0.084	-0.166
	(0.160)	(0.280)	(0.384)	(0.355)
Panel D: Privatization: initial state-owned employment share				
Regional tariff reduction (RTR)	-0.090	-0.490***	-1.235***	-1.580***
	(0.134)	(0.163)	(0.172)	(0.205)
Quartile indicators, 1995 state-owned	✓	✓	✓	✓
employment share distirbution				
Panel E: Privatization: change in state-owned employment share, 1995 to t				
Regional tariff reduction (RTR)	-0.096	-0.514***	-1.243***	-1.558***
	(0.120)	(0.149)	(0.146)	(0.182)
Change in state-owned employment share		0.095	0.286	0.176
		(0.178)	(0.214)	(0.227)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. 475 microregion observations. Regional earnings premia calculated controlling for age, sex, education, and industry of employment. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium. See text for detailed description of each panel and for control construction. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.8.4   Commodity Price Boom

Figures B4 - B6 show price indexes for major Brazilian export commodities from 1991 to 2010, using data from the Primary Commodity Price Series produced by the International Monetary Fund. Although there is some variability across commodities, there is little evidence of the commodity price boom prior to 2004. For example, in Figure B4, aggregate agricultural commodity prices were nearly identical in 1991 and 2003, and Brazil's main agricultural exports (coffee, sugar, soy, and cotton) all have lower average prices in 2003 than in 1991. Natural resource and meat commodities in Figures B5 and B6 exhibit similar patterns, with relatively flat series prior to 2003 and substantial growth for many commodities starting in 2004.

   Note the contrast between this time series pattern and the effects of regional tariff reductions on formal earnings and employment growth in Figures 3 and 4. The earnings effects grow steadily from 1996 to 2003, in spite of the fact that most commodity prices actually fell a bit during that time span. Commodity prices start growing very quickly in 2004 and later, during which the earnings effects start to level off. A similar argument applies to the employment effects, which grow from 1994 to 2004 and level off subsequently. Thus, the timing of the commodity price boom does not conform with the timing of the earnings and employment effects, making it very unlikely that commodity prices drive our results, particularly before 2004.

   To reinforce this time-series evidence, in Table B9 we implement a wide variety of tests to rule out the commodity price boom as a potential confounder. Panel A reproduces the main specification for comparison. Panels B and C respectively restrict the sample of regions to those with below median and bottom quartile employment shares in agriculture and mining, the sectors affected by the commodity price boom. Note that mining includes fuel extraction. When focusing on regions with minimal exposure to commodity sectors, we find even larger growth in the effects of liberalization on earnings than in the entire sample. In Panel D, we maintain all regions, but only calculate earnings premia for workers employed in the manufacturing sector, omitting workers in commodity and nontradables sectors. Once again, the earnings effects continue to grow substantially over time given this restriction. As an aside, note that earnings in the manufacturing sector, which most directly experienced the effects of trade liberalization, exhibit significant effects on impact, in 1995. This finding suggests that the *very* short-run effects of liberalization were concentrated in the industries facing the largest tariff cuts, but that the earnings effects spread out to other sectors over time through labor market equilibrium.

   In Panels E and F, rather than restricting the sample of regions or workers, we control for the commodity price boom directly. We utilize the regional commodity price shocks constructed by Adão (2015). Special thanks to Rodrigo Adão for sharing his data and code. See Appendix C in Adão (2015) for details on the data source and his equation (16) for the shock construction. To summarize, he calculates commodity-specific changes in log price from 1991 to each subsequent year using data from the Commodity Research Bureau, and constructs regional weighted averages of these commodity price shocks. The weights reflect each commodity sector's share of total labor payments in all commodity sectors in the region in 1991. Because this measure does not incorporate regional variation in overall exposure to commodity price changes, we also flexibly control for the regional importance of commodities by including quartile indicators for the region's 1991 share of regional employment in agriculture and mining. As seen in Panel E, controlling directly for these commodity price movements has little influence on the increasing profile of earnings effects, either before or after the beginning of commodity price boom.

We implement a similar exercise in Panel F of Table B9 using more detailed commodity price information from the IMF Primary Commodity Price Series, a subset of which is shown in Figures B4 - B6. While Adão's measure uses only 6 aggregate commodity indexes, our alternative measure uses 19 separate indexes. As an example, this detail allows us to distinguish between the commodity price changes faced by regions specialized in coffee vs. sugar, which are grouped together under Adão's classification. Table B10 shows our mapping from commodity industry codes in the 1991 Census to the IMF price indexes. We calculate the change in log price index from 1991 to each subsequent year for each IMF commodity and then generate regional weighted averages of these price changes, where weights reflect the relevant commodity's share of regional employment in 1991. As seen in Panel F, including these commodity price controls has no substantive effect upon the relationship between regional tariff reductions and regional earnings, either before or after the beginning of commodity price boom. Because our regional commodity price shock measure already incorporates the overall share of commodity industries in regional employment, we do not additionally control for that share, though doing so has no substantive effect on the results.

The rise of China appears to have played a substantial role in driving up commodity prices in the late 2000s. As a final test of the commodity price boom hypothesis, we follow Costa et al. (2016), who study the regional effects of import competition from China and Chinese demand for exports. Rather than focusing on commodity *prices*, Costa et al. study the effects of import and export *quantity* shocks, along the lines of Autor et al. (2013). They construct industry-level Chinese import supply (IS) and export demand (XD) shocks as the growth in industry imports from or exports to China from 2000 to 2010, divided by Brazilian employment in the industry in 2000. They then generate regional weighted average shocks using the year 2000 industry distribution of employment in each region. Finally, they instrument for these shocks using similar measures based on the growth in Chinese trade to countries other than Brazil. Special thanks to Francisco Costa for providing us with their shock and instrument measures.

Because Costa et al. examine shocks and outcomes between 2000 and 2010, in Panel A of Table B11 we provide our baseline earnings estimate for this time period, with a base year of 2000 rather than 1991. We use a slightly more aggregate region definition to match theirs, yielding 405 region observations. The coefficient estimate of -1.068 in column (1) is nearly identical to the difference between the estimates for 2000 and 2010 in columns (3) and (6), respectively, of Panel A in Table 2 of -1.065. The slight difference results from the difference in region definitions between the two tables. In columns (2)-(4) of Table B11, we introduce the Chinese import supply and export demand shocks, instrumented following Costa et al.. The two shocks have the expected sign, with increased import competition lowering regional earnings and increased export demand increasing them (very slightly), though only the import supply shock is statistically different from zero. This result might seem surprising, given that Costa et al. find significant effects of export demand on wages. However, when they control for the regional composition of workers and for outcome pre-trends, as we do here, the export result loses statistical significance (see their Table 2, Panel B, column (5) in their paper). When we include these controls, they have only a very small effect on our coefficient of interest, further confirming that the divergence in earnings growth between regions facing larger and smaller tariff reductions was not driven by China's effects on commodity markets.

Figure B4: Agricultural Commodity Prices - 1991-2010
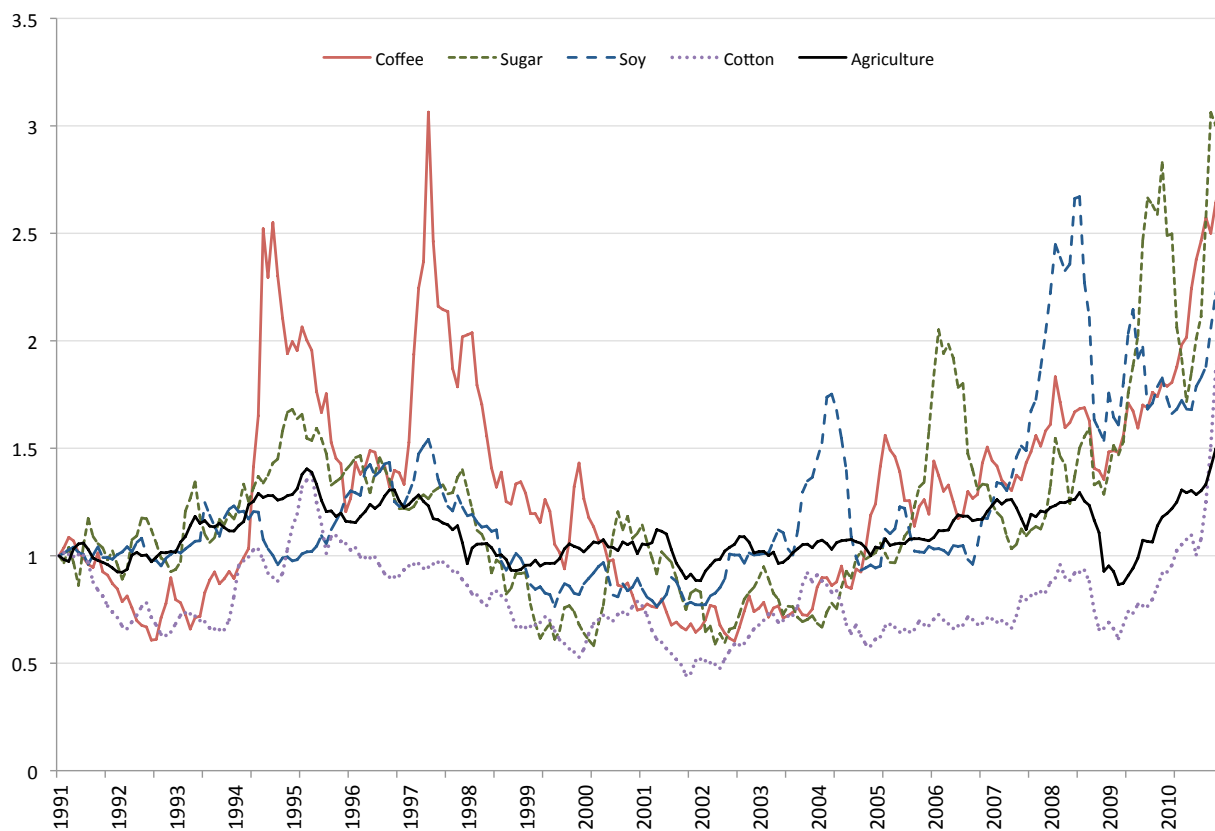
Monthly price series from IMF Primary Commodity Price Series, rescaled to equal 100 in January 1991.
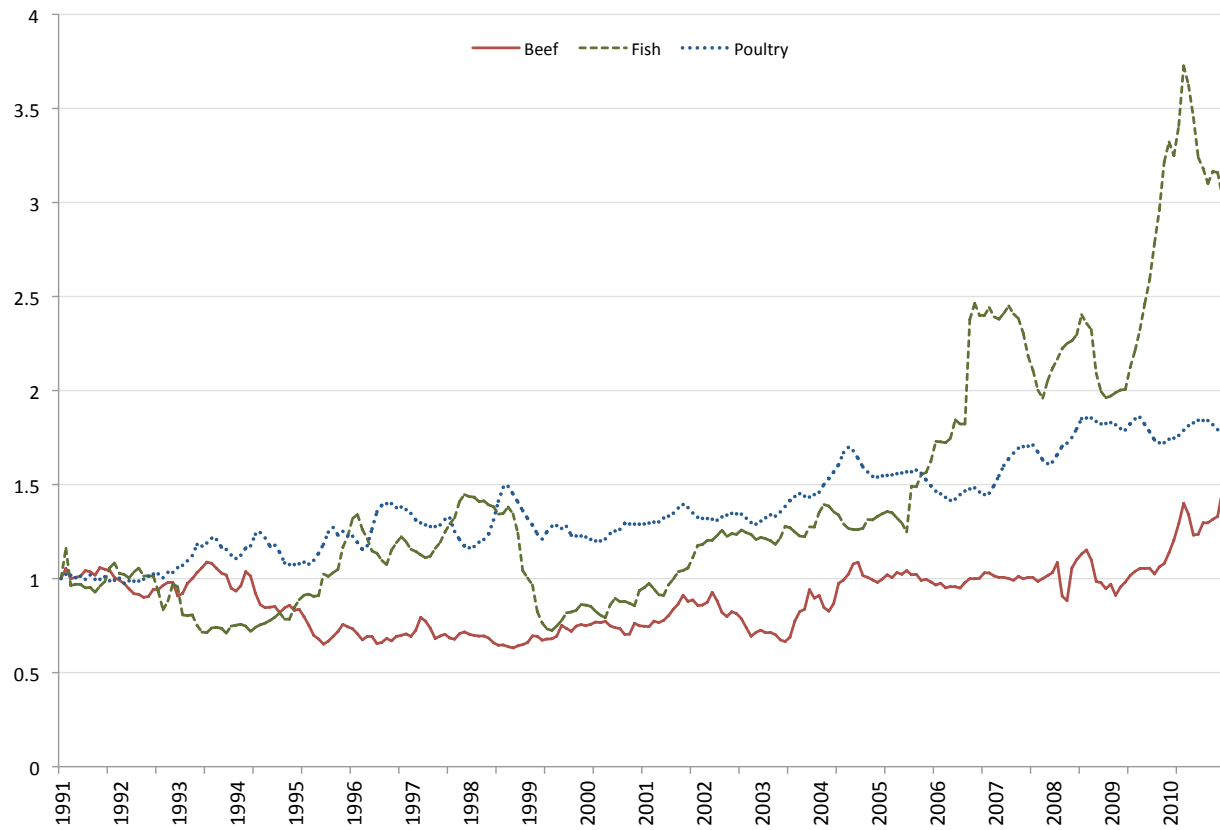
Figure B5: Resource Commodity Prices - 1991-2010
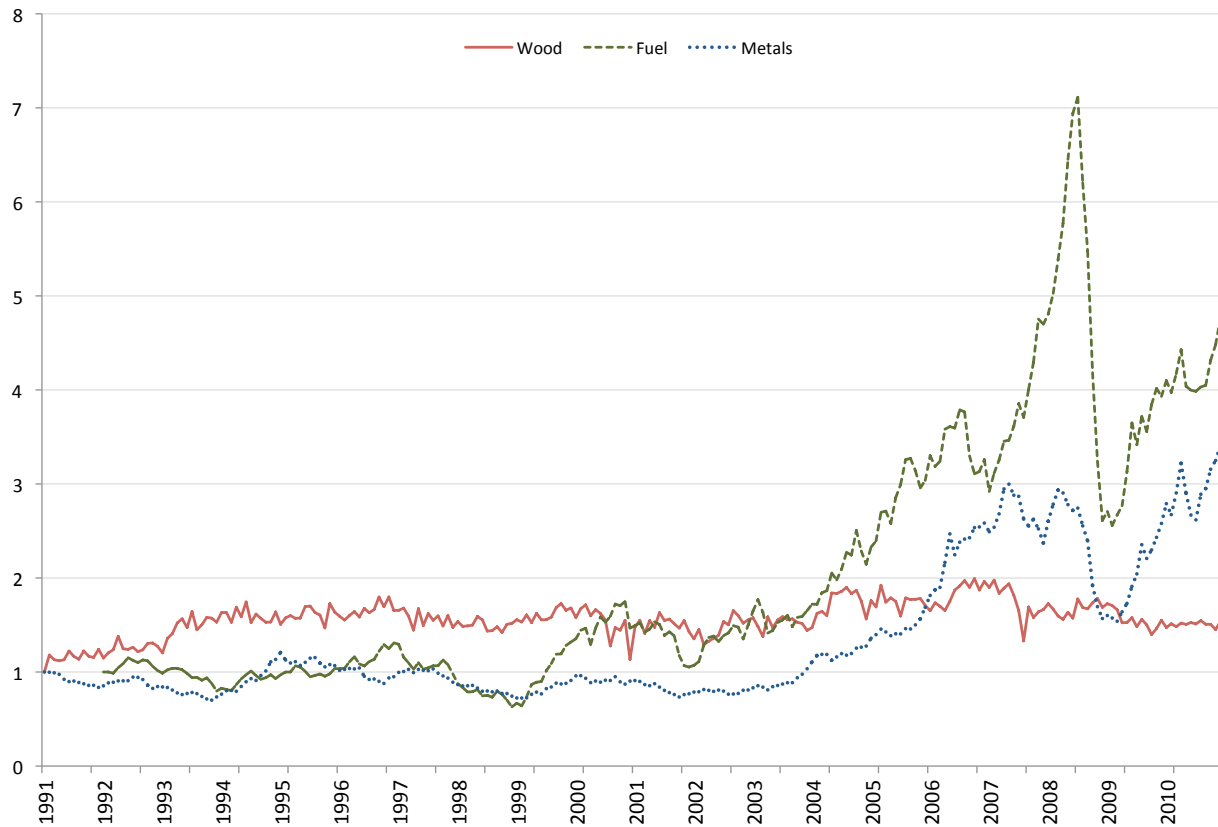
Monthly price series from IMF Primary Commodity Price Series, rescaled to equal 100 in January 1991, except the fuel index which begins in 1992 and is rescaled to 1 in January 1992.

Figure B6: Meat Commodity Prices - 1991-2010

Monthly price series from IMF Primary Commodity Price Series, rescaled to equal 100 in January 1991.

Table B9: Commodity Price Boom: Regional log Formal Earnings Premia - 1995, 2000, 2005, 2010

Change in log Formal Earnings Premia:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Main specification				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Panel B: Below-median agriculture/mining employment share (238 obs)				
Regional tariff reduction (RTR)	-0.017	-0.534***	-1.424***	-1.829***
	(0.148)	(0.165)	(0.163)	(0.213)
Panel C: Bottom quartile agriculture/mining employment share 118 obs)				
Regional tariff reduction (RTR)	0.006	-0.347	-1.379***	-2.153***
	(0.267)	(0.262)	(0.280)	(0.373)
Panel D: Manufacturing sector earnings				
Regional tariff reduction (RTR)	-0.501***	-0.965***	-1.878***	-2.252***
	(0.158)	(0.192)	(0.214)	(0.262)
Panel E: Direct commodity price controls per Adao (2015)				
Regional tariff reduction (RTR)	-0.052	-0.290	-1.269***	-1.926***
	(0.259)	(0.257)	(0.276)	(0.372)
Regional commodity price shocks	0.033	-0.039	0.118	0.045
(Adao 2015)	(0.207)	(0.167)	(0.092)	(0.127)
Quartile indicators, 1991 agriculture/mining employment share distribution	✓	✓	✓	✓
Panel F: Direct commodity price controls using detailed commodity price data (IMF)				
Regional tariff reduction (RTR)	0.023	-0.591***	-1.210***	-1.718***
	(0.143)	(0.188)	(0.137)	(0.330)
Regional commodity price shocks	0.179	0.160	0.421	-0.069
(IMF data)	(0.120)	(0.266)	(0.277)	(0.149)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. 475 microregion observations unless otherwise noted (Panels B and C). Regional earnings premia calculated controlling for age, sex, education, and industry of employment. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium. See text for detailed description of each panel and for control construction. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

Table B10: Mapping from Commodity Industries to IMF Price Indexes

1991 Census Industry (*atividade*)		IMF Index	
11	Agave	PRAWM	Agricultural Raw Materials
12	Cotton	PCOTTIND	Cotton
13	Rice	PRICENPQ	Rice
14	Banana	PBANSOP	Bananas
15	Cocoa	PCOCO	Cocoa beans
16	Coffee	PCOFFOTM	Coffee
17	Sugar cane	PSUGAISA	Sugar
18	Tobacco	PRAWM	Agricultural Raw Materials
19	Cassava	PRAWM	Agricultural Raw Materials
20	Corn	PMAIZMT	Maize (corn)
21	Soybeans	PSOYB	Soybeans
22	Wheat	PWHEAMT	Wheat
23	Horticulture and floriculture	PRAWM	Agricultural Raw Materials
24	Forestry	PSAWORE	Soft Sawnwood
25	Other agricultural products	PRAWM	Agricultural Raw Materials
26	Livestock	PBEEF	Beef
27	Aviculture	PPOULT	Poultry (chicken)
28	Beekeeping and Sericulture	PRAWM	Agricultural Raw Materials
29	Other livestock	PRAWM	Agricultural Raw Materials
31	Rubber	PRUBB	Rubber
32	Yerba mate	PRAWM	Agricultural Raw Materials
33	Plant fibres	PRAWM	Agricultural Raw Materials
34	Fruits, oilseeds, and waxes	PRAWM	Agricultural Raw Materials
35	Wood	PSAWORE	Soft Sawnwood
36	Charoal	PCOALAU	Coal
37	Other harvesting activities	PRAWM	Agricultural Raw Materials
41	Fishing	PFISH	Fishmeal
42	Aquaculture	PFISH	Fishmeal
51	Oil and natural gas mining	PNRG	Fuel (Energy)
52	Coal mining	PCOALAU	Coal
55	Metallic mineral panning and deposition	PMETA	Metals
56	Radioactive mineral mining	PURAN	Uranium
58	Metallic mineral mining (except those in other categories)	PMETA	Metals
581	Agriculture and livestock services	PRAWM	Agricultural Raw Materials

Table B11: Regional log Formal Earnings Premia with Costa et al (2015) controls

Change in log Formal Earnings Premia, 2000-2010:	(1) OLS	(2) IV	(3) IV	(4) IV
Regional tariff reduction (RTR)	-1.068***	-0.929***	-1.069***	-0.931***
	(0.111)	(0.123)	(0.107)	(0.122)
Formal earnings pre-trend (86-90)	-0.077	-0.064	-0.076	-0.063
	(0.053)	(0.055)	(0.051)	(0.055)
China import supply (Costa et al. 2015)		-0.034***		-0.034***
		(0.010)		(0.010)
China export demand (Costa et al. 2015)			0.001	0.001
			(0.002)	(0.002)
State fixed effects (26)	✓	✓	✓	✓
R-squared	0.733	0.738	0.733	0.739
1st stage F (Kleibergen-Paap)		22.16	441.6	11.31

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. See text for description of China import supply and export demand controls and associated instruments from Costa et al. (2015). First stage partial F-statistics reported in brackets. 405 microregion observations. Regional earnings premia calculated controlling for age, sex, education, and industry of employment. Standard errors (in parentheses) adjusted for 90 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.9   Earnings and Employment Sample Splits

Tables B12 and B13 present earnings results splitting the sample of workers into those employed in the tradable sector (Panel B), those employed in the nontradable sector (Panel C), more educated (Panel D), and less educated (Panel E). Note that earnings and employment effects grow for all subsamples. The employment effects are concentrated in the tradable sector and among less skilled workers, though panels D and E in both tables should be interpreted with care, as the regional tariff reduction shocks are derived from a model with a single type of labor. For a more general model with two skill types, see Dix-Carneiro and Kovak (2015a).

Table B12: Sample Splits: Regional log Formal Earnings Premia - 1995, 2000, 2005, 2010

Change in log Formal Earnings Premia:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Full sample				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Panel B: Tradable sector workers				
Regional tariff reduction (RTR)	-0.287*	-0.754***	-1.623***	-1.934***
	(0.149)	(0.184)	(0.203)	(0.254)
Panel C: Nontradable sector workers				
Regional tariff reduction (RTR)	-0.060	-0.389**	-1.143***	-1.401***
	(0.150)	(0.179)	(0.169)	(0.183)
Panel D: More educated workers (high school or more)				
Regional tariff reduction (RTR)	0.310*	-0.539***	-1.611***	-2.053***
	(0.160)	(0.173)	(0.192)	(0.224)
Panel E: Less educated workers (less than high school)				
Regional tariff reduction (RTR)	-0.218*	-0.626***	-1.354***	-1.758***
	(0.121)	(0.143)	(0.137)	(0.180)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. 475 microregion observations. Regional earnings premia calculated controlling for age, sex, education, and industry of employment. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

Table B13: Sample Splits: Regional log Formal Employment - 1995, 2000, 2005, 2010

Change in log Formal Employment:	1991-1995	1991-2000	1991-2005	1991-2010
	(1)	(2)	(3)	(4)
Panel A: Full sample				
Regional tariff reduction (RTR)	-1.900***	-3.533***	-4.517***	-4.663***
	(0.422)	(0.582)	(0.685)	(0.679)
Panel B: Tradable sector workers				
Regional tariff reduction (RTR)	-5.790***	-8.416***	-10.097***	-10.156***
	(0.850)	(0.993)	(1.101)	(1.140)
Panel C: Nontradable sector workers				
Regional tariff reduction (RTR)	0.726	-0.733	-1.500*	-1.600**
	(0.455)	(0.664)	(0.778)	(0.749)
Panel D: More educated workers (high school or more)				
Regional tariff reduction (RTR)	0.141	-1.219*	-1.637**	-2.195***
	(0.450)	(0.644)	(0.772)	(0.733)
Panel E: Less educated workers (less than high school)				
Regional tariff reduction (RTR)	-2.507***	-4.556***	-6.328***	-6.910***
	(0.465)	(0.626)	(0.770)	(0.787)
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal employment in regions facing larger tariff reductions. 475 microregion observations. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

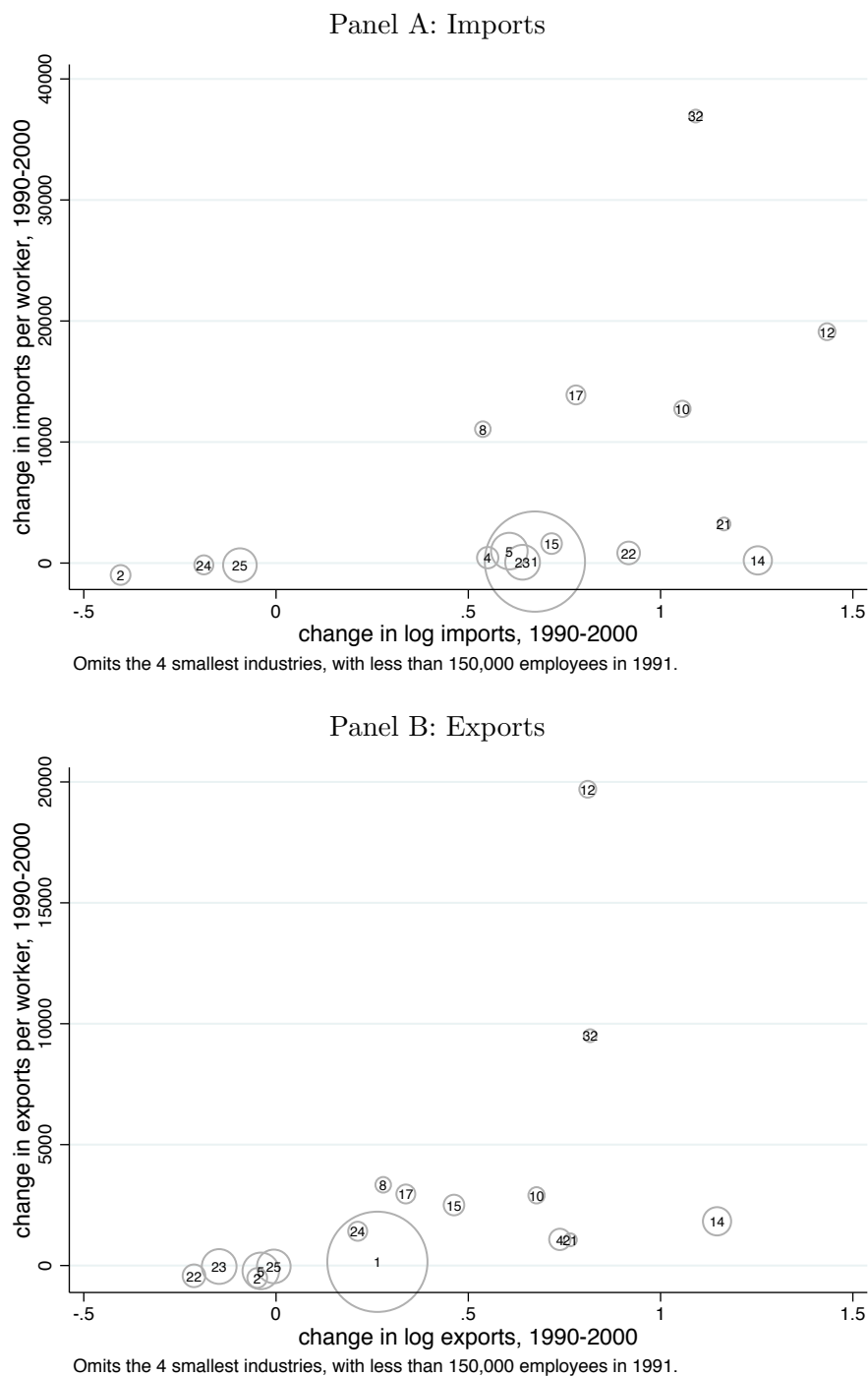## B.10   Regional Change in log Imports and Exports

This section presents versions of Figure 5 and Table 6 using alternative trade quantity measures reflecting regional weighted averages of the change in log imports or exports rather than the change in imports or exports per worker. These alternative measures are presented only for descriptive purposes and as a statistical robustness test.

   We use the change in trade per worker in the main text both because it is theoretically justified, as shown by Autor et al. (2013), and because it intuitively captures the effects of changing trade quantities on local labor market outcomes. Figures B7 and B8, which show scatter plots relating the industry-level change in trade per worker and the change in log trade for imports and exports in 2000 and 2010. Markers are proportional to 1991 employment, and industry labels correspond to Table A1. These plots show that the change in log trade often deviates substantially from the change in trade per worker. This deviation occurs primarily in industries with relatively small trade flows and relatively large employment. As an example, consider the Wood, Furniture, and Peat industry (code 14) in Panel A of Figure B7. In this industry, Brazil imported R\$71 million in 1990 and R\$249 million in 2000 (all values in year 2005 Reais). This very large proportional growth in imports corresponds to the large value for the change in log imports of 1.25. However, initial employment in this industry was also quite large, 822,579, so the change in imports per worker was only R\$216, much smaller than the values in the thousands or tens of thousands in other industries. Therefore, although the amount of imports increased very much in proportional terms, it was still insignificant compared to the number of workers in the industry. The change in trade per worker captures the relative scale of trade and employment, while the change in log trade does not.

   Nonetheless, figure B9 shows the relationships between RTR_r and the regional change in log imports and exports (paralleling Figure 5). See (18) and (19) in Appendix A.6 for details on constructing the change in log trade measures. Regions facing larger tariff reductions experience larger increases in log imports and larger declines in log exports. The magnitude of each effect grows over time, suggesting that perhaps slow trade quantity responses could explain the slow growth of regional earnings and employment effects in Figures 3. We demonstrate that this is not the case by directly controlling for the regional change in log import and export measures when estimating the effect of RTR_r on regional earnings growth. The specifications in Table B14 parallel those in Table 6, using the alternative change-in-log measures of trade flows for both the controls and the instruments. In all cases, the earnings effects of liberalization grow even more when controlling for import and export quantity growth than in the baseline specification in Panel A. The relevant Stock and Yogo (2005) critical value for the first-stage F-statistic is 21, so Panel C exhibits a potential weak instruments problem. We therefore present two additional sets of results in Table B15, using the change in trade per worker measure when calculating the instruments rather than the change in log trade flow measure. The weak instrument issue is not longer present, and the effects of RTR_r on regional earnings still increase substantially over time.

   As with the standard trade-per-worker trade flow measures considered in the main text, the analysis using the alternative change-in-log trade flow measure rules out slow import or export responses as the mechanism driving the slowly growing earnings effects. This is to be expected given the discussion at the beginning of this section, since the change in log trade measure does not well capture the effects of changing trade flows on workers' labor market outcomes.
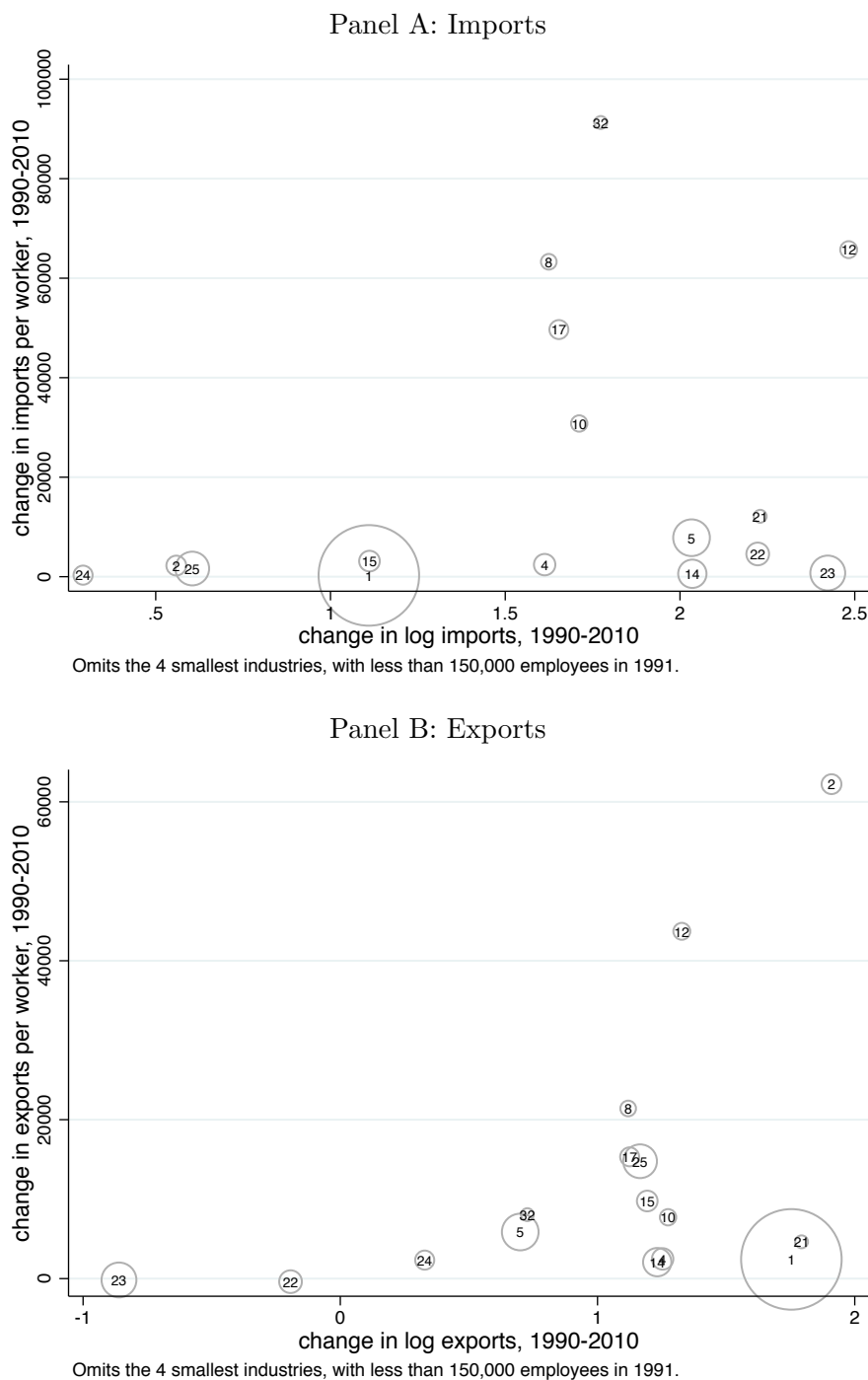
Figure B7: Change in Trade Per Worker vs. Change in log Trade - 1990-2000

Panel A: Imports

Omits the 4 smallest industries, with less than 150,000 employees in 1991.

Panel B: Exports

Omits the 4 smallest industries, with less than 150,000 employees in 1991.

Each point is an industry, with labels corresponding to Table A1. The y-axis measures the change in trade per worker initially employed in the industry (in 1991) and the x-axis measures the change in log trade. The figure omits the 4 smallest industries in terms of 1991 employment, which often fall well outside the scale shown. Because they are small, they receive very little weight in the regional analysis that forms this paper's main analysis. Marker size proportional to 1991 employment.

Figure B8: Change in Trade Per Worker vs. Change in log Trade - 1990-2010

Panel A: Imports

Omits the 4 smallest industries, with less than 150,000 employees in 1991.

Panel B: Exports

Omits the 4 smallest industries, with less than 150,000 employees in 1991.

Each point is an industry, with labels corresponding to Table A1. The y-axis measures the change in trade per worker initially employed in the industry (in 1991) and the x-axis measures the change in log trade. The figure omits the 4 smallest industries in terms of 1991 employment, which often fall well outside the scale shown. Because they are small, they receive very little weight in the regional analysis that forms this paper's main analysis. Marker size proportional to 1991 employment.

Figure B9: Regional Imports and Exports, Change in log Measure - 1992-2010
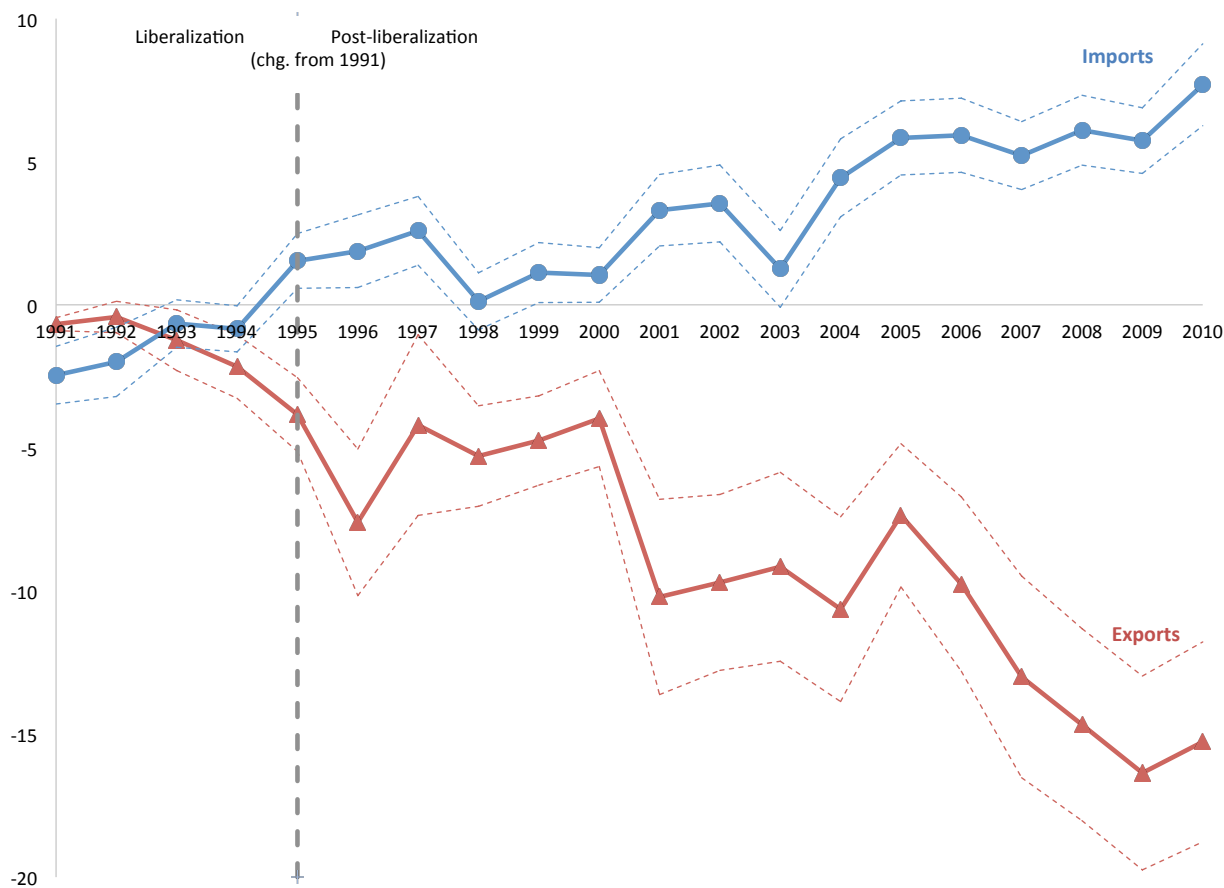
Each point reflects an individual regression coefficient, $\hat{\theta}_t$, following (3), where the dependent variable is the change in regional imports (blue circles) or exports using the change in log measures described in (18) and (19) in Appendix A.6. The independent variable is the regional tariff reduction (RTR), defined in (2). Note that the RTR always reflects tariff reductions from 1990-1995. All regressions include state fixed effects, but do *not* include pre-liberalization trends due to a lack of Comtrade trade data before 1989. Positive (negative) estimates imply larger increases in trade flow in regions facing larger (smaller) tariff reductions. Vertical bar indicates that liberalization was complete by 1995. Dashed lines show 95 percent confidence intervals. Standard errors adjusted for 112 mesoregion clusters.

Table B14: Slow Response of Imports or Exports, Change in log Measure - 1995, 2000, 2005, 2010 (Part 1 of 2)

Change in log Formal Earnings Premia:	1991-1995 (1)	1991-2000 (2)	1991-2005 (3)	1991-2010 (4)
Panel A: Main specification				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Panel B: Controls for trade quantitities (OLS)				
Regional tariff reduction (RTR)	-0.142	-0.581***	-1.476***	-1.887***
	(0.124)	(0.142)	(0.151)	(0.212)
Import quantity control (change in log)		0.025**		
		(0.012)		
Export quantity control (change in log)		-0.009		
		(0.008)		
Panel C: Latin America IV				
Regional tariff reduction (RTR)	-0.236	-0.608***	-1.479***	-2.158***
	(0.145)	(0.150)	(0.343)	(0.579)
Import quantity control (change in log)		-0.009		
		(0.062)		
Export quantity control (change in log)		-0.052		
		(0.039)		
First-stage F (Kleibergen-Paap)		19.71		
Panel D: Colombia IV				
Regional tariff reduction (RTR)	-0.227	-0.570***	-1.316***	-1.985***
	(0.146)	(0.160)	(0.306)	(0.509)
Import quantity control (change in log)		-0.044		
		(0.049)		
Export quantity control (change in log)		-0.059*		
		(0.032)		
First-stage F (Kleibergen-Paap)		61.45		
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

Negative coefficient estimates for the regional tariff reduction (RTR) imply larger declines in formal earnings in regions facing larger tariff reductions. Panel A replicates the earnings results in columns (3) and (6) of Table 2 and in Figure 3. Panels B-D control for regional import and export quantities, calculated using the change in log trade flows; see Appendix A.6 for details. These panels stack the data across years, allowing the effect of RTR_r to vary over time but fixing the import and export quantity coefficients over time. We instrument for the potentially endogenous import and export controls using regional measures of commodity price growth from Adão (2015) and with regional trade flows for other countries. We consider the combination of Argentina, Chile, Colombia, Paraguay, Peru, and Uruguay ("Latin America") Colombia alone. In each case, we measure imports and exports between these countries and the rest of the world, excluding Brazil. This gives us 2 endogenous variables and 57 instruments (= 3 instruments × 19 years). First-stage Kleinbergen-Paap F statistics are shown, for comparison to the Stock and Yogo (2005) critical value of 21 to reject 5 percent bias relative to OLS. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. Efficiency weighted by the inverse of the squared standard error of the estimated change in log formal earnings premium. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

Table B15: Slow Response of Imports or Exports, Change in log Measure - 1995, 2000, 2005, 2010 (Part 2 of 2)

Change in log Formal Earnings Premia:	1991-1995 (1)	1991-2000 (2)	1991-2005 (3)	1991-2010 (4)
Panel E: Latin America IV (change in trade per worker)				
Regional tariff reduction (RTR)	-0.063	-0.487***	-1.143***	-1.364**
	(0.147)	(0.156)	(0.333)	(0.561)
Import quantity control (change in log)		-0.022		
		(0.050)		
Export quantity control (change in log)		0.005		
		(0.031)		
First-stage F (Kleibergen-Paap)		31.44		
Panel F: Colombia IV (change in trade per worker)				
Regional tariff reduction (RTR)	-0.096	-0.529***	-1.294***	-1.594***
	(0.120)	(0.141)	(0.139)	(0.169)
Import quantity control (change in log)		-0.045		
		(0.055)		
Export quantity control (change in log)		-0.026		
		(0.026)		
First-stage F (Kleibergen-Paap)		61.58		
Formal earnings pre-trend (86-90)	✓	✓	✓	✓
State fixed effects (26)	✓	✓	✓	✓

See Table B14 for general notes. Because the instrument in Panel C of Table B14 was marginally weak, Panels E and F present versions using instruments based on the change in imports per worker, while the trade quantity controls are calculated using the change in log. Both specifications reject the weak instrument concern.

B.11   Overall Employment

Table B16 shows the effect of liberalization on overall regional employment, including both formally and informally employed workers. We use Census data to capture informally employed individuals, and control for 1980-1991 and 1970-1980 outcome pre-trends. The estimates vary substantially across specifications and all but one are insignificantly different from zero. These results provide little evidence in favor of overall employment as a potential source of agglomeration economies.

Table B16: Regional log Overall Employment - 2000, 2010

Change in log Overall Employment:	1991-2000			1991-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Regional tariff reduction (RTR)	0.203	-0.419	-0.272	0.657**	-0.478	-0.265
	(0.209)	(0.450)	(0.260)	(0.314)	(0.683)	(0.410)
Overall employment pre-trend (80-91)	0.329**		0.281**	0.538***		0.454**
	(0.136)		(0.130)	(0.202)		(0.189)
Overall employment pre-trend (70-80)		0.221***	0.120***		0.385***	0.230***
		(0.071)	(0.035)		(0.093)	(0.052)
State fixed effects (26)	✓	✓	✓	✓	✓	✓
R-squared	0.563	0.479	0.585	0.574	0.484	0.609

Positive (negative) coefficient estimates for the regional tariff reduction imply larger increases (decreases) in overall employment in regions facing larger tariff reductions. Outcomes calculated using Census data. 405 microregion observations. Efficiency weighted by the inverse of the squared standard error of the dependent variable estimate. Pre-trends computed for 1980-1991 and 1970-1980. Standard errors (in parentheses) adjusted for 112 mesoregion clusters. *** Significant at the 1 percent, ** 5 percent, * 10 percent level.

B.12 Capital Adjustment Confidence Intervals

Figures B10 - B12 show the capital adjustment profiles in Figure 6, including 95-percent confidence intervals, which were omitted from Figure 6 for clarity.

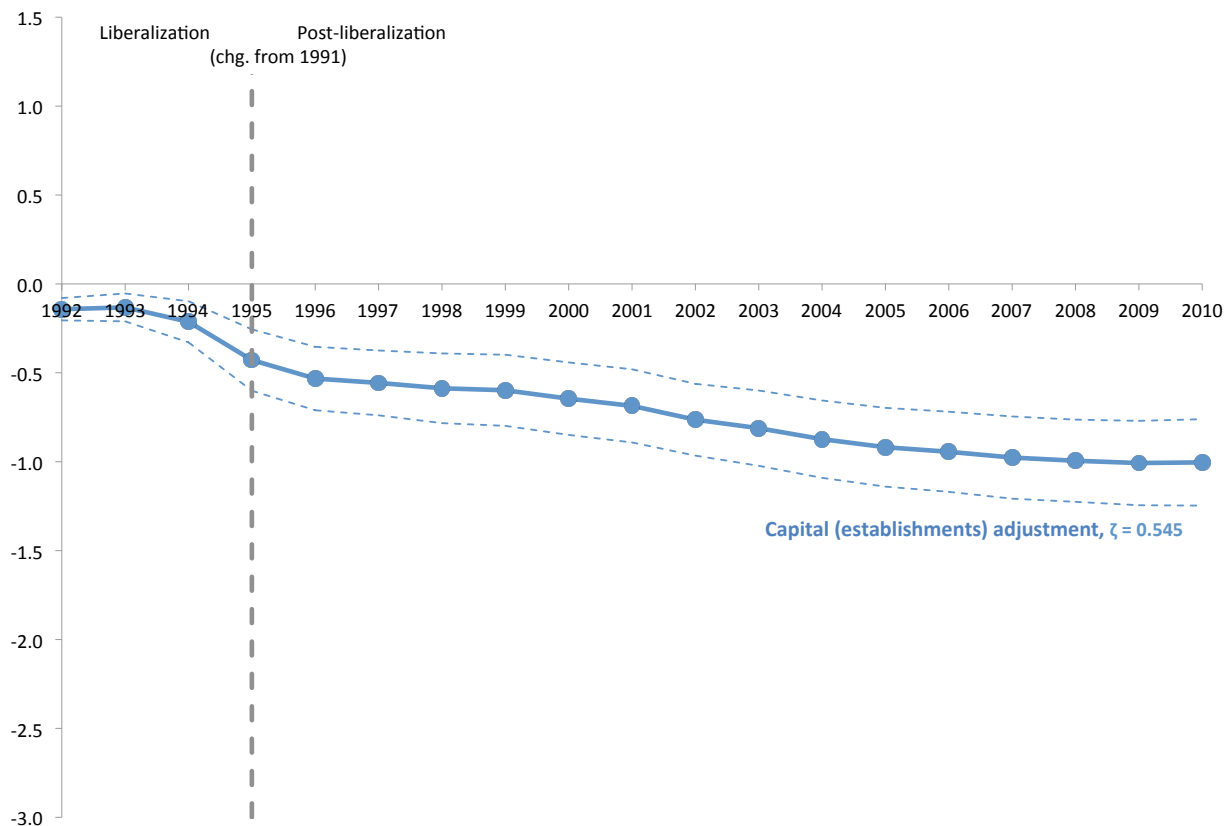Figure B10: Capital Adjustment Quantification - Low ζ - 1992-2010
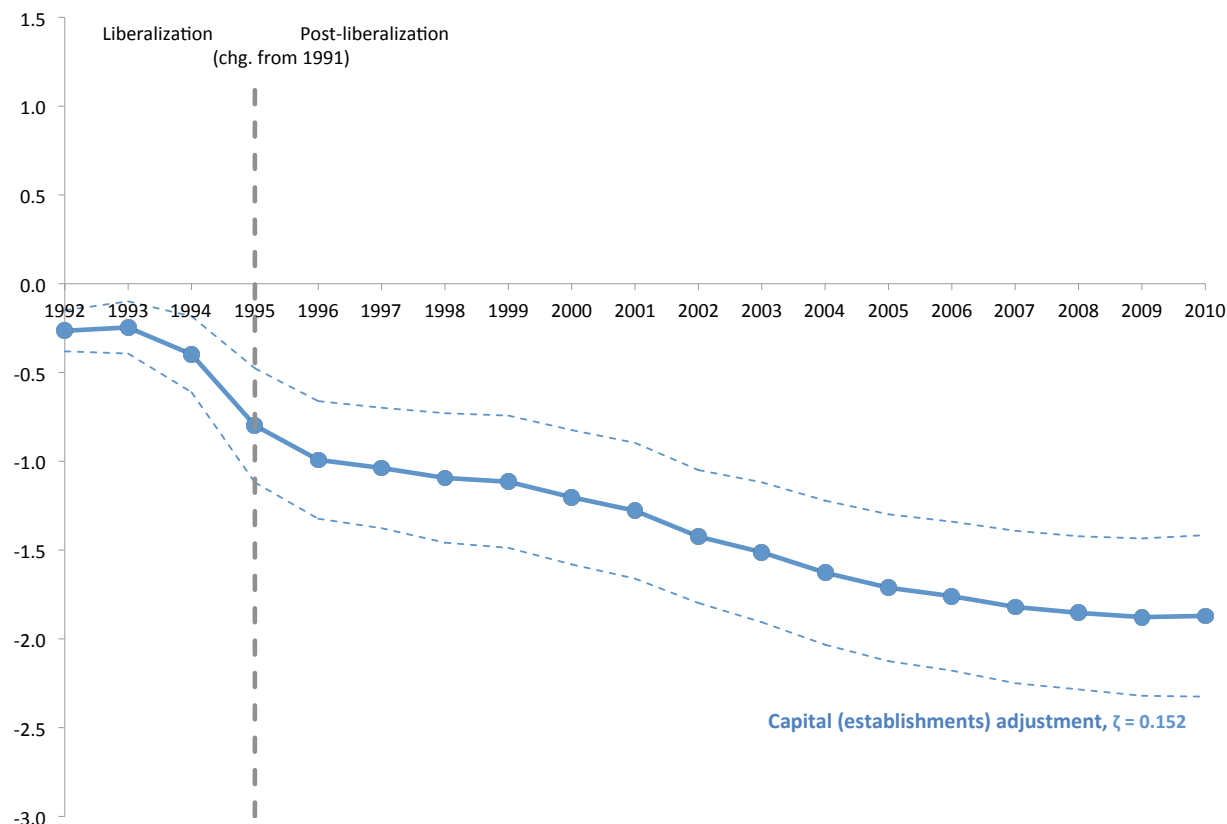
Each point reflects an individual regression coefficient, $\hat{\theta}_t$, following (3). The dependent variable is capital's contribution to overall adjustment, using the number of regional formal establishments as a proxy for regional capital. This figure shows the profile using the low estimate of $\zeta = 0.152$. The independent variable is the regional tariff reduction (RTR), defined in (2). Note that the RTR always reflects tariff reductions from 1990-1995. All regressions include state fixed effects, and post-liberalization regressions control for the 1986-1990 outcome pre-trend. Negative estimates imply larger declines in the number of establishments in regions facing larger tariff reductions. Vertical bar indicates that liberalization was complete by 1995. Dashed lines show 95 percent confidence intervals. Standard errors adjusted for 112 mesoregion clusters.

Figure B11: Capital Adjustment Quantification - Mid ζ - 1992-2010
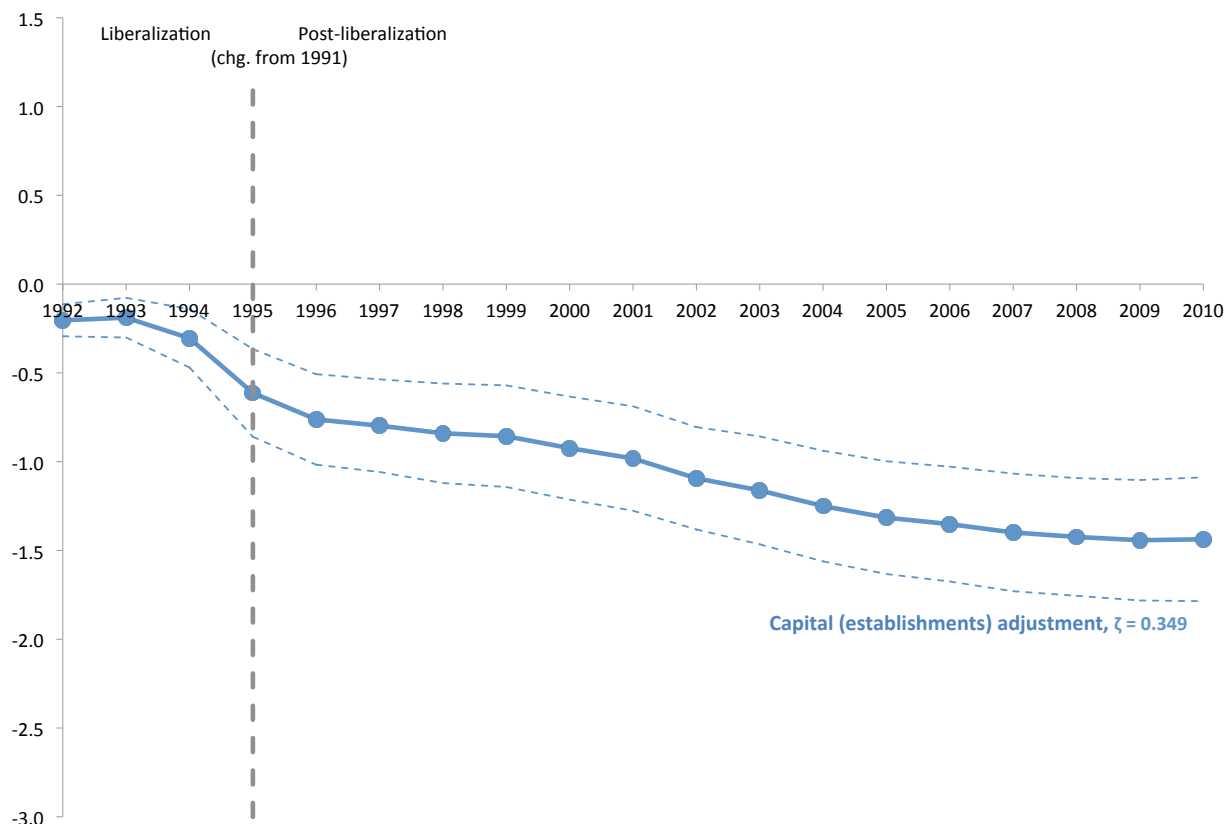
Each point reflects an individual regression coefficient, $\hat{\theta}_t$, following (3). The dependent variable is capital's contribution to overall adjustment, using the number of regional formal establishments as a proxy for regional capital. This figure shows the profile using the low estimate of $\zeta = 0.349$. The independent variable is the regional tariff reduction (RTR), defined in (2). Note that the RTR always reflects tariff reductions from 1990-1995. All regressions include state fixed effects, and post-liberalization regressions control for the 1986-1990 outcome pre-trend. Negative estimates imply larger declines in the number of establishments in regions facing larger tariff reductions. Vertical bar indicates that liberalization was complete by 1995. Dashed lines show 95 percent confidence intervals. Standard errors adjusted for 112 mesoregion clusters.

Figure B12: Capital Adjustment Quantification - High ζ - 1992-2010

Each point reflects an individual regression coefficient, $\hat{\theta}_t$, following (3). The dependent variable is capital's contribution to overall adjustment, using the number of regional formal establishments as a proxy for regional capital. This figure shows the profile using the low estimate of $\zeta = 0.545$. The independent variable is the regional tariff reduction (RTR), defined in (2). Note that the RTR always reflects tariff reductions from 1990-1995. All regressions include state fixed effects, and post-liberalization regressions control for the 1986-1990 outcome pre-trend. Negative estimates imply larger declines in the number of establishments in regions facing larger tariff reductions. Vertical bar indicates that liberalization was complete by 1995. Dashed lines show 95 percent confidence intervals. Standard errors adjusted for 112 mesoregion clusters.

B.13 Exit by Establishment Size

Here we examine the relationship between establishment exit and RTR_r, separately by initial establishment size. We run the following specification at the establishment-year level, using the sample of all active establishments in 1991.

$$Exit_{irt} = \sum_{k=1}^{6} \beta_t^k Size_i^k \cdot RTR_r + \sum_{k=1}^{6} \phi_k Size_i^k + \gamma_t NT_i + \vartheta_t PreExit_r^{1986-1990} + \epsilon_{irt} \qquad (26)$$

where $Size_i^k$ is an indicator for whether establishment i fell into size bin k in 1991, NT_i is an indicator for establishments in the nontradable sector, and $PreExit_r^{1986-1990}$ is a pre-trend control for the share of regional establishments in 1986 that shut down between 1986 and 1990.

Figure B13 plots the β_t^k coefficients, with the relevant initial employment bin definitions shown on the right side. Although there is some variation across establishment sizes, with more exit among larger establishments than smaller establishments, it is clear that exit rates increased throughout the size distribution for establishments whose regions faced larger tariff declines.

Figure B13: Regional log Cumulative Formal Establishment Exit, by Establishment Size - 1992-2010
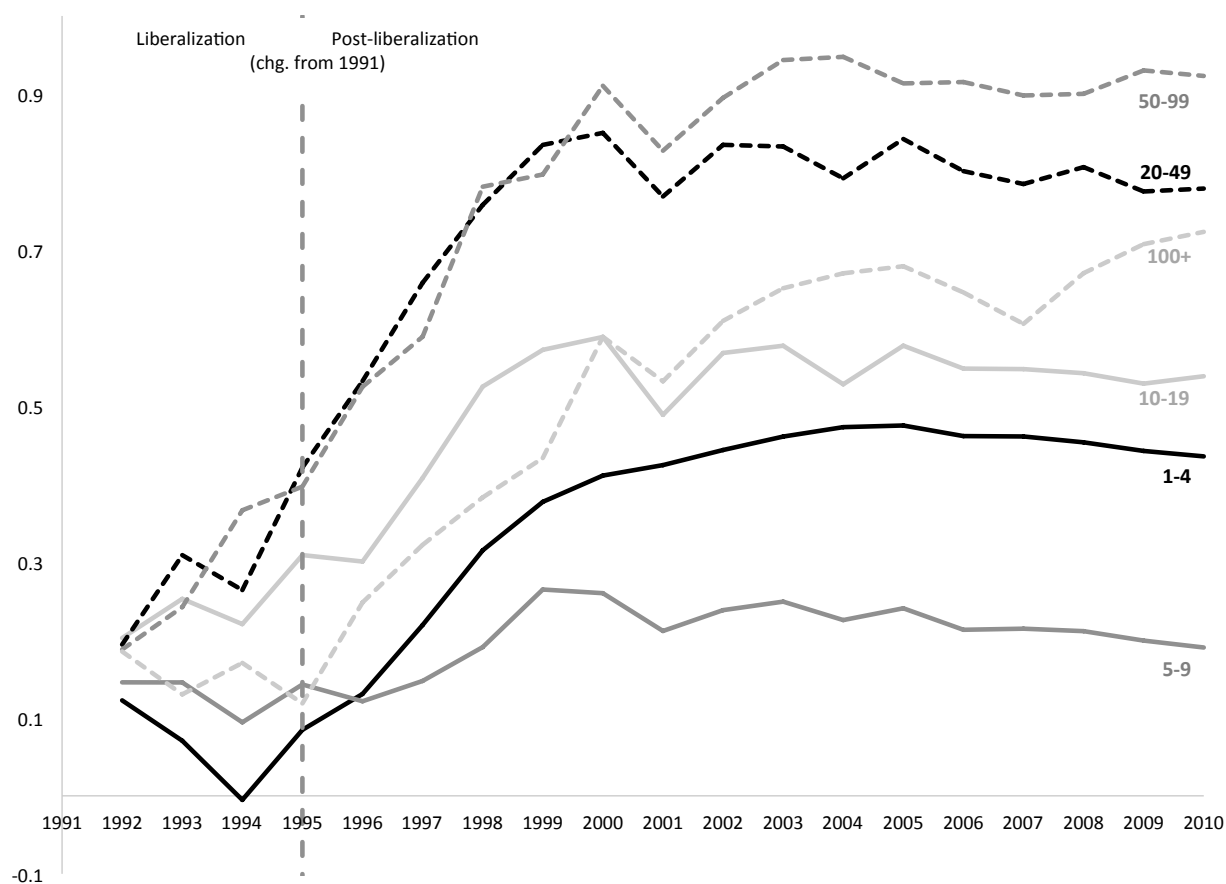
Plots the β_t^k coefficients in (26), estimated using the sample of all active establishments in 1991. The size range, indexed by k, is reported at the right side of each profile. Vertical bar indicates that liberalization was complete by 1995.

C   Model

C.1   Baseline Model

This section generalizes the specific-factors model of regional economies from Kovak (2013) to allow for changes in regional productivity (agglomeration economies) and changes in labor and capital inputs.

The economy consists of many regions, indexed by r, which may produce goods in many industries, indexed by i. Production in each industry uses Cobb-Douglas technology with constant returns to scale and three inputs: labor, a fixed industry-specific factor, and capital. Labor, L_r, is assumed to be perfectly mobile between industries within a region. The industry-specific factor, T_{ri}, is usable only in its respective region and industry and is fixed over time. Capital, K_{ri}, is usable only in its respective region and industry but may change over time. Output of industry i in region r is

$$Y_{ri} = A_{ri} L_{ri}^{1-\varphi_i} \left(T_{ri}^{\zeta_i} K_{ri}^{1-\zeta_i} \right)^{\varphi_i}, \tag{C1}$$

where $\varphi_i, \zeta_i \in (0,1)$. To allow for the possibility of agglomeration economies and factor adjustment, we allow A_{ri}, L_r, and K_{ri} to change over time. Goods and factor markets are perfectly competitive, and producers face exogenous prices P_i, common across regions and fixed by world prices and tariffs.

Consider a particular region r, and suppress the region subscript. Let a_{Li}, a_{Ti}, and a_{Ki} be the respective amounts of labor, specific factor, and capital used in producing one unit of Y_i. Regional factor market clearing implies

$$\sum_i a_{Li} Y_i = L, \tag{C2}$$

$$a_{Ti} Y_i = T_i \quad \forall i, \tag{C3}$$

$$a_{Ki} Y_i = K_i \quad \forall i. \tag{C4}$$

Perfect competition implies that the price equals factor payments,

$$a_{Li} w + a_{Ti} s_i + a_{Ki} R_i = P_i \quad \forall i, \tag{C5}$$

where w is the wage, s_i is the specific-factor price, and R_i is the price of capital. Define $\hat{x}$ as the proportional change in x, and differentiate (C5).

$$(1 - \varphi_i)\hat{w} + \varphi_i \zeta_i \hat{s}_i + \varphi_i(1-\zeta_i)\hat{R}_i = \hat{P}_i + \hat{A}_i \quad \forall i, \tag{C6}$$

which uses the fact that, from cost minimization,

$$(1 - \varphi_i)\hat{a}_{Li} + \varphi_i \zeta_i \hat{a}_{Ti} + \varphi_i(1-\zeta_i)\hat{a}_{Ki} = -\hat{A}_i \quad \forall i. \tag{C7}$$

Differentiate the factor market clearing conditions.

$$\sum_i \lambda_i(\hat{a}_{Li} + \hat{Y}_i) = \hat{L}, \tag{C8}$$

$$\hat{Y}_i = -\hat{a}_{Ti} \quad \forall i, \tag{C9}$$

$$\hat{Y}_i = \hat{K}_i - \hat{a}_{Ki} \quad \forall i, \tag{C10}$$

where $\lambda_i \equiv \frac{L_i}{L}$ is the share of regional labor allocated to industry i, and we use the fact that $\hat{T}_i = 0$. With Cobb-Douglas production, the elasticity of substitution is one, so

$$\hat{a}_{Ki} - \hat{a}_{Ti} = \hat{s}_i - \hat{R}_i, \tag{C11}$$

$$\hat{a}_{Li} - \hat{a}_{Ki} = \hat{R}_i - \hat{w}, \tag{C12}$$

Combining (C8), (C10), and (C12) yields

$$\sum_i \lambda_i \hat{R}_i - \hat{w} = \hat{L} - \sum_i \lambda_i \hat{K}_i. \tag{C13}$$

Combine (C9), (C10), and (C11) to yield

$$\hat{s}_i = \hat{R}_i + \hat{K}_i \quad \forall i. \tag{C14}$$

Plug this into (C6) and simplify.

$$\hat{R}_i = \frac{\hat{P}_i + \hat{A}_i - (1 - \varphi_i)\hat{w} - \varphi_i \zeta_i \hat{K}_i}{\varphi_i} \tag{C15}$$

Finally, plug this into (C13), solve for $\hat{w}$, and restore regional subscripts to yield the equilibrium relationship for regional wage changes, equation (9) in the main text.

$$\hat{w}_r = \sum_i \beta_{ri} \hat{P}_i + \sum_i \beta_{ri} \hat{A}_{ri} - \delta_r \left(\hat{L}_r - \sum_i \lambda_{ri}(1 - \zeta_i)\hat{K}_{ri} \right) \tag{C16}$$

$$\text{where} \quad \beta_{ri} \equiv \frac{\lambda_{ri}\frac{1}{\varphi_i}}{\sum_j \lambda_{rj}\frac{1}{\varphi_j}} \quad \text{and} \quad \delta_r \equiv \frac{1}{\sum_j \lambda_{rj}\frac{1}{\varphi_j}}.$$

C.2   Agglomeration Economies

As discussed in the main text, when examining agglomeration economies and quantifying the long-run effects of slow capital adjustment and agglomeration, we assume perfectly mobile capital in the long run ($\hat{R}_r = \hat{R} \ \forall r$), and identical technology across industries ($\varphi_i = \varphi \ \forall i$ and $\zeta_i = \zeta \ \forall i$). The assumption of perfectly mobile capital allows us to substitute out the change in capital, $\hat{K}_{ri}$ for the change in its price, $\hat{R}$, which is constant across industries and regions.

Start with the labor market clearing condition in (C8), substitute in the specific-factors clearing condition in (C9) and the Cobb-Douglas conditions in (C11) and (C12) to yield

$$\sum_i \lambda_i \hat{s}_i - \hat{w} = \hat{L}. \tag{C17}$$

Rearrange the zero-profit condition in (C6) to solve for $\hat{s}_i$,

$$\hat{s}_i = \frac{1}{\varphi\zeta}(\hat{P}_i + \hat{A}_i) - \frac{\varphi(1 - \zeta)}{\varphi\zeta}\hat{R} - \frac{1 - \varphi}{\varphi\zeta}\hat{w}, \tag{C18}$$

and plug it into (C17). Solving for $\hat{w}$ and restoring regional subscripts yields the following expression.

$$\hat{w}_r = \frac{1}{1 - \varphi(1-\zeta)} \sum_i \beta_{ri}(\hat{P}_i + \hat{A}_{ri}) - \frac{\varphi\zeta}{1 - \varphi(1-\zeta)}\hat{L}_r - \frac{\varphi(1-\zeta)}{1 - \varphi(1-\zeta)}\hat{R} \tag{C19}$$

$$\text{where} \quad \beta_{ri} \equiv \frac{\lambda_{ri}\frac{1}{\varphi}}{\sum_j \lambda_{rj}\frac{1}{\varphi}} = \lambda_{Lri}$$

To incorporate agglomeration economies, we assume a constant elasticity agglomeration function, (11), and a constant labor supply elasticity, (12). Substituting these into (C19) and simplifying yields the following expression for the regional wage change, equation (13) in the main text, which we use to estimate the agglomeration elasticity, κ.

$$\hat{w}_r = \frac{\eta}{\eta[1 - \varphi(1-\zeta)] - \kappa + \varphi\zeta} \sum_i \beta_{ri}\hat{P}_i - \frac{\varphi(1-\zeta)\eta}{\eta[1 - \varphi(1-\zeta)] - \kappa + \varphi\zeta}\hat{R} \tag{C20}$$

We also use an alternative employment-based approach to estimate κ. Start by noting that employment in a region $\times$ industry pair is given by $L_{ri} \equiv a_{Lri}Y_{ri}$. Differentiating this definition and plugging in the specific-factor market clearing condition, (C9), and the Cobb-Douglas substitution conditions, (C11) and (C12), we have

$$\hat{L}_i = \hat{s}_i - \hat{w}. \tag{C21}$$

Substitute in $\hat{s}_i$ from (C18) and simplify.

$$\hat{L}_i = \frac{1}{\varphi\zeta}(\hat{P}_i + \hat{A}_i) - \frac{1 - \varphi(1-\zeta)}{\varphi\zeta}\hat{w} - \frac{\varphi(1-\zeta)}{\varphi\zeta}\hat{R} \tag{C22}$$

Plug in the labor supply and agglomeration equations, (12) and (11).

$$\hat{L}_i = \frac{1}{\varphi\zeta}\hat{P}_i - \frac{\eta[1 - \varphi(1-\zeta)] - \kappa}{\eta\varphi\zeta}\hat{w} - \frac{\varphi(1-\zeta)}{\varphi\zeta}\hat{R} \tag{C23}$$

Finally, plug in the equilibrium wage change in (C20), combine terms, and restore regional subscripts to yield equation (14) in the main text.

$$\hat{L}_{ri} = \frac{1}{\varphi\zeta}\hat{P}_i - \frac{1}{\varphi\zeta} \cdot \frac{\eta[1 - \varphi(1-\zeta)] - \kappa}{\eta[1 - \varphi(1-\zeta)] - \kappa + \varphi\zeta} \sum_i \beta_{ri}\hat{P}_i - \frac{\varphi(1-\zeta)}{\eta[1 - \varphi(1-\zeta)] - \kappa + \varphi\zeta}\hat{R} \tag{C24}$$