Online appendixes for "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" by Brian K. Kovak

Specific Factors Model

A1. Factor prices

This section closely follows Jones (1975), but deviates from that paper's result by allowing the amount of labor available to the regional economy to vary. Consider a particular region, r, suppressing that subscript on all terms. Industries are indexed by $i = 1...N$. L is the total amount of labor and T_i is the amount of industry i-specific factor available in the region. a_{Li} and a_{Ti} are the respective quantities of labor and specific factor used in producing one unit of industry i output. Letting Y_i be the output in each industry, the factor market clearing conditions are

(A1) $$a_{Ti}Y_i = T_i \qquad \forall i,$$

(A2) $$\sum_i a_{Li}Y_i = L.$$

Under perfect competition, the output price equals the factor payments, where w is the wage and R_i is the specific factor price.

(A3) $$a_{Li}w + a_{Ti}R_i = P_i \qquad \forall i$$

Let hats represent proportional changes, and consider the effect of price changes $\hat{P}_i$. θ_i is the cost share of the specific factor in industry i.

(A4) $$(1-\theta_i)\hat{w} + \theta_i\hat{R}_i = \hat{P}_i \qquad \forall i,$$

which follows from the envelope theorem result that unit cost minimization implies

(A5) $$(1-\theta_i)\hat{a}_{Li} + \theta_i\hat{a}_{Ti} = 0 \qquad \forall i.$$

Differentiate (A1), keeping in mind that T_i is fixed in all industries.

(A6) $$\hat{Y}_i = -\hat{a}_{Ti} \qquad \forall i$$

Similarly, differentiate (A2), let $\lambda_i = \frac{L_i}{L}$ be the fraction of regional labor utilized in industry i, and substitute in (A6) to yield

(A7) $$\sum_i \lambda_i(\hat{a}_{Li} - \hat{a}_{Ti}) = \hat{L}.$$

By the definition of the elasticity of substitution between T_i and L_i in production,

$$\hat{a}_{Ti} - \hat{a}_{Li} = \sigma_i(\hat{w} - \hat{R}_i) \qquad \forall i. \tag{A8}$$

Substituting this into (A7) yields

$$\sum_i \lambda_i \sigma_i (\hat{R}_i - \hat{w}) = \hat{L}. \tag{A9}$$

Equations (A4) and (A9) can be written in matrix form as follows.

$$\left[\begin{array}{cccc|c} \theta_1 & 0 & \dots & 0 & 1-\theta_1 \\ 0 & \theta_2 & \dots & 0 & 1-\theta_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \theta_N & 1-\theta_N \\ \hline \lambda_1\sigma_1 & \lambda_2\sigma_2 & \dots & \lambda_N\sigma_N & -\sum_i \lambda_i\sigma_i \end{array}\right] \left[\begin{array}{c} \hat{R}_1 \\ \hat{R}_2 \\ \vdots \\ \hat{R}_N \\ \hline \hat{w} \end{array}\right] = \left[\begin{array}{c} \hat{P}_1 \\ \hat{P}_2 \\ \vdots \\ \hat{P}_N \\ \hline \hat{L} \end{array}\right] \tag{A10}$$

Rewrite this expression as follows for convenience of notation.

$$\left[\begin{array}{c|c} \mathbf{\Theta} & \boldsymbol{\theta}_{\mathbf{L}} \\ \hline \boldsymbol{\lambda}' & -\sum_i \lambda_i\sigma_i \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{R}} \\ \hline \hat{w} \end{array}\right] = \left[\begin{array}{c} \hat{\mathbf{P}} \\ \hline \hat{L} \end{array}\right] \tag{A11}$$

Solve for $\hat{w}$ using Cramer's rule and the rule for the determinant of partitioned matrices.

$$\hat{w} = \frac{\hat{L} - \boldsymbol{\lambda}'\mathbf{\Theta}^{-1}\hat{\mathbf{P}}}{-\sum_i \lambda_i\sigma_i - \boldsymbol{\lambda}'\mathbf{\Theta}^{-1}\boldsymbol{\theta}_{\mathbf{L}}} \tag{A12}$$

Note that the inverse of the diagonal matrix $\mathbf{\Theta}$ is a diagonal matrix of $\frac{1}{\theta_i}$'s. This yields the effect of goods price changes and changes in regional labor on regional wages:

$$\hat{w} = \frac{-\hat{L}}{\sum_{i'} \lambda_{i'} \frac{\sigma_{i'}}{\theta_{i'}}} + \sum_i \beta_i \hat{P}_i \tag{A13}$$

$$\text{where} \qquad \beta_i = \frac{\lambda_i \frac{\sigma_i}{\theta_i}}{\sum_{i'} \lambda_{i'} \frac{\sigma_{i'}}{\theta_{i'}}} \tag{A14}$$

This expression with $\hat{L} = 0$ yields (1). Changes in specific factor prices can be calculated from wage changes by rearranging (A4).

$$\hat{R}_i = \frac{\hat{P}_i - (1-\theta_i)\hat{w}}{\theta_i} \tag{A15}$$

Plugging in (A13) and collecting terms yields the effect of goods price changes and changes in regional labor on specific factor price changes.

$$\text{(A16)} \qquad \hat{R}_i = \frac{(1-\theta_i)}{\theta_i} \frac{\hat{L}}{\sum_{i'} \lambda_{i'} \frac{\sigma_{i'}}{\theta_{i'}}} + \left(\beta_i + \frac{1}{\theta_i}(1-\beta_i)\right)\hat{P}_i - \frac{(1-\theta_i)}{\theta_i}\sum_{k\neq i}\beta_k \hat{P}_k$$

Setting $\hat{L} = 0$ in (A13) and (A16) yields the equivalent expressions in Jones (1975).

A2. Graphical representation of the model

The equilibrium adjustment mechanisms at work in the model are demonstrated graphically in Figure A1, which represents a two-region ($r = 1, 2$) and two-industry ($i = A, B$) version of the model.[29] Region 1 is relatively well endowed with the industry A specific factor. In each panel, the x-axis represents the total amount of labor in the country to be allocated across the two industries in the two regions, and the y-axis measures the wage in each region. Focusing on the left portion of panel (a), the curve labeled $P_A F_L^A$ is the marginal value product of labor in industry A, and the curve labeled $P_B F_L^B$ is the marginal value product of labor in industry B, measuring the amount of labor in industry B from right to left. Given labor mobility across sectors, the intersection of the two marginal value product curves determines the equilibrium wage, and the allocation of labor in region 1 between industries A and B, as indicated on the x-axis. The right portion of panel (a) is interpreted similarly for region 2. For visual clarity, the figures were generated assuming equal wages across regions before any price changes. This assumption is not necessary for any of the theoretical results presented in the paper.

Panel (a) of Figure A1 shows an equilibrium in which wages are equalized across regions. Since region 1 is relatively well endowed with industry A specific factor, it allocates a greater share of its labor to industry A. Panel (b) shows the effect of a 50% decrease in the price of good A, so the good A marginal value product curve in both regions moves down halfway toward the x-axis. Consistent with (1), the impact of this price decline is greater in region 1, which allocated a larger fraction of labor to industry A than did region 2. Thus, region 1's wage falls more than region 2's wage. Now workers in region 1 have an incentive to migrate to region 2. For each worker that migrates, the central vertical axis moves one unit to the left, indicating that there are fewer laborers in region 1 and more in region 2. As the central axis shifts left, so do the two marginal value product curves that are measured with respect to that axis. This shift raises the wage in region 1 and lowers the wage in region 2. Panel (c) shows the resulting equilibrium

[29]Figure A1 was generated under the following conditions. Production is Cobb-Douglas with specific-factor cost share equal to 0.5 in both industries. $\bar{L} = 10$, $T_{1A} = 1$, $T_{1B} = 0.4$, $T_{2A} = 0.4$, and $T_{2B} = 1$. Initially, $P_A = P_B = 1$, and after the price change, $P_A = 0.5$.

assuming costless migration and equalized wage across regions. Again, none of the theoretical results in the paper require this assumption.

A3. Nontraded goods prices

As above, consider a particular region, omitting the r subscript on all terms. Industries are indexed by $i = 1...N$. The final industry, indexed N, is nontraded, while other industries ($i \neq N$) are traded. The addition of a nontraded industry does not alter the results of the previous section, but makes it necessary to describe regional consumers' preferences to fix the nontraded good's equilibrium price.

Assume a representative consumer with Cobb-Douglas preferences over goods from each industry. This implies the following relationship between quantity demanded by consumers, Y_i^c, total consumer income, m, and price, P_i.

$$\hat{Y}_i^c = \hat{m} - \hat{P}_i \tag{A17}$$

Consumers own all factors and receive all revenue generated in the economy; $m = \sum_i P_i Y_i^p$ where Y_i^p is the amount of good i produced in equilibrium. Applying the envelope theorem to the revenue function for the regional economy, one can show[30]

$$\hat{m} = \eta_L \hat{L} + \sum_i \varphi_i \hat{P}_i. \tag{A18}$$

where η_L is the share of total factor payments accounted for by wages and φ_i is the share of regional production value accounted for by industry i. Plugging (A18) into (A17) for the nontraded industry and rearranging terms,

$$\hat{P}_N - \sum_{i \neq N} \frac{\varphi_i}{\sum_{i' \neq N} \varphi_{i'}} \hat{P}_i - \frac{\eta_L}{1 - \varphi_N} \hat{L} = \frac{-1}{1 - \varphi_N} \hat{Y}_N^c. \tag{A19}$$

Holding labor fixed, this expression shows that consumption shifts *away* from the nontraded good if the nontraded price increases relative to a weighted average of traded goods prices, with weights based on each good's share of traded sector output value.

A similar expression can be derived for the production side of the model. Wage equals the value of marginal product, so $\hat{w} = \hat{P}_N + \hat{F}_L^N$. The specific factor is fixed, so $\hat{Y}_N^p = (1 - \theta_N)\hat{L}_N$. Combining these two observations with Euler's theorem

[30] The revenue maximization problem is specified as follows.

$$r(P_1...P_N, L) = \max_{L_i} \sum_i P_i F^i(L_i, T_i) \quad s.t. \quad \sum_i L_i \leq L$$

and the definition of the elasticity of substitution as in footnote 6 yields

$$\hat{w} = \hat{P}_N - \frac{\theta_N}{\sigma_N(1-\theta_N)}\hat{Y}_N^p \tag{A20}$$

Substitute in the expression for $\hat{w}$ in (A13) and rearrange.

$$\hat{P}_N - \sum_{i \neq N} \frac{\beta_i}{\sum_{i' \neq N} \beta_{i'}} \hat{P}_i + \frac{1}{(1-\beta_N)\sum_{i'} \lambda_{i'} \frac{\sigma_{i'}}{\theta_{i'}}} \hat{L} = \frac{\theta_N}{(1-\beta_N)\sigma_N(1-\theta_N)} \hat{Y}_N^p \tag{A21}$$

Holding labor fixed, this expression shows that production shifts *toward* the nontraded good if the nontraded price increases relative to a weighted average of traded goods prices, with weights based on the industry's size and labor demand elasticity captured in β_i.

Equations (A19) and (A21) relate very closely to the intuition described in Section I.B for the case of only one traded good. Consumption shifts away from the nontraded good if its price rises relative to average traded goods prices while production shifts toward it. Since regional consumption and production of the nontraded good must be equal, with a single traded good the nontraded price must exactly track the traded good price. Combining (A19) and (A21) shows that in the many-traded-good case, the nontraded price change is a weighted average of traded goods price changes.

$$\hat{P}_N = \frac{\eta_L - \frac{\sigma_N}{\theta_N} \frac{(1-\theta_N)}{\sum_i \lambda_i \frac{\sigma_i}{\theta_i}}}{\sum_{i' \neq N} \left[\frac{\sigma_N}{\theta_N}(1-\theta_N)\beta_{i'} + \varphi_{i'} \right]} \hat{L} + \sum_{i \neq N} \xi_i \hat{P}_i \tag{A22}$$

$$\text{where} \qquad \xi_i = \frac{\frac{\sigma_N}{\theta_N}(1-\theta_N)\beta_i + \varphi_i}{\sum_{i' \neq N} \left[\frac{\sigma_N}{\theta_N}(1-\theta_N)\beta_{i'} + \varphi_{i'} \right]} \tag{A23}$$

A4. Restriction to Drop the Nontraded Sector from Weighted Averages

Under Cobb-Douglas production with equal factor shares across industries, $\theta_i = \theta$, and $\sigma_i = 1\ \forall i$. This implies that $\beta_i = \lambda_i$, the fraction of labor in each industry. Similarly,

$$\varphi_i \equiv \frac{P_i Y_i^p}{\sum_{i'} P_{i'} Y_{i'}^p} = \frac{\lambda_i \eta_L}{1-\theta} = \lambda_i. \tag{A24}$$

since $\eta_L = 1 - \theta$ when $\theta_i = \theta$. When $\beta_i = \varphi_i$, the weights ξ_i in (A22) are

$$\xi_i = \frac{\beta_i}{\sum_{i' \neq N} \beta_{i'}} \tag{A25}$$

Plug these weights into (3) and then plugging that into (1) yields a result equivalent to dropping the nontraded sector from the weighted average in (1) and (2).

$$\hat{w} = \frac{\sum_{i \neq N} \beta_i \hat{P}_i}{\sum_{i' \neq N} \beta'_i} \tag{A26}$$

A5. Wage Impact of Changes in Regional Labor

This section shows that an increase in regional labor decreases the regional wage. Substituting (A22) into (A13), holding traded goods prices fixed ($\hat{P}_i = 0 \; \forall i \neq N$), and rearranging yields

$$\hat{w} = \left[\frac{-\frac{\sigma_N}{\theta_N}(1-\theta_N) - (1-\varphi_N) + \lambda_N \frac{\sigma_N}{\theta_N}\eta_L}{\left(\sum_i \lambda_i \frac{\sigma_i}{\theta_i}\right) \left[\frac{\sigma_N}{\theta_N}(1-\theta_N)(1-\beta_N) + (1-\varphi_N)\right]} \right] \hat{L}. \tag{A27}$$

Since the denominator is strictly positive, the sign of the relationship between $\hat{L}$ and $\hat{w}$ is determined by the numerator. An increase in regional labor will decrease the wage if an only if

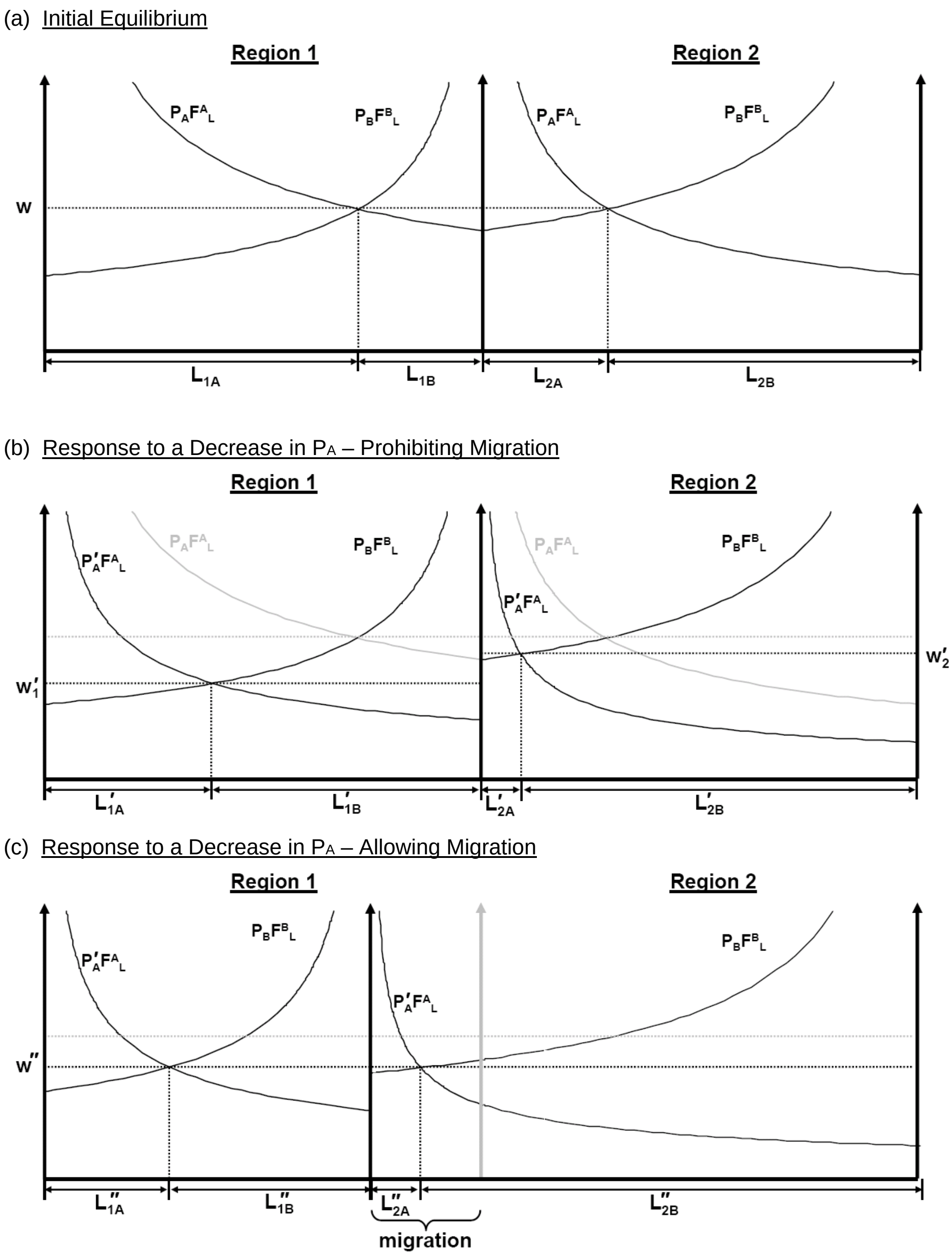
$$(1-\varphi_N) > \frac{\sigma_N}{\theta_N}(\lambda_N \eta_L - (1-\theta_N)). \tag{A28}$$

Using the fact that $\eta_L = (\sum_i \lambda_i (1-\theta_i)^{-1})^{-1}$ the previous expression is equivalent to

$$(1-\varphi_N) > \frac{\sigma_N}{\theta_N}(1-\theta_N)\left(\frac{\lambda_N(1-\theta_N)^{-1}}{\sum_i \lambda_i (1-\theta_i)^{-1}} - 1\right). \tag{A29}$$

The left hand side is positive while the right hand side is negative. Thus, the inequality always holds, and an increase in regional labor will always lower the regional wage.

Figure A1. Graphical Representation of Specific Factors Model of Regional Economies



Data Appendix

B1. Industry Crosswalk

National accounts data from IBGE and trade policy data from (Kume, Piani and de Souza 2003) are available by industry using the the Nível 50 and Nível 80 classifications. The 1991 Census reports individuals' industry of employment using the *atividade* classification system, while the 2000 Census reports industries based on the newer *Classificação Nacional de Atividades Econômicas - Domiciliar* (*CNAE-Dom*). Table B1 shows the industry definition used in this paper and its concordance with the various other industry definitions in the underlying data sources. The concordance for the 1991 Census is based on a crosswalk between the national accounts and *atividade* industrial codes published by IBGE. The concordance for the 2000 Census is based on a crosswalk produced by IBGE's Comissão Nacional de Classificação, available at *http://www.ibge.gov.br/concla/*.

B2. Trade Policy Data

Nível 50 trade policy data come from Kume, Piani and de Souza (2003). Depending on the time period, they aggregated tariffs on 8,750 - 13,767 individual goods first using unweighted averages to aggregate individual goods up to the Nível 80 level, and then using value added weights to aggregate from Nível 80 to Nível 50.[31] In order to maintain this weighting scheme, I weight by value added when aggregating from Nível 50 to the final classification listed in Table B1. For reference, Figure B1 and Figure B2 show the evolution of nominal tariffs and effective rates of protection in the ten largest sectors by value added. Along with the general reduction in the level of protection, the dispersion in protection was also greatly reduced, consistent with the goal of aligning domestic production incentives with world prices. It is clear that the move from a high-level, high-dispersion tariff structure to a low-level low-dispersion tariff distribution generated substantial variation in protection changes across industries; industries with initially high levels of protection experienced the largest cuts, while those with initially lower levels experienced smaller cuts.

Recall from Section III that the 1987-1990 period exhibited substantial tariff redundancy due to the presence of selective import bans and special import regimes that exempted many imports from the full tariff charge. The tariff changes during this period did not generate a change in the protection faced by producers, but rather reflected a process called called *tarifação*, or "tariffization," in which the import bans and special import regimes were replaced by tariffs providing equivalent levels of protection (Kume, Piani and de Souza 2003). The nominal tariff increases in 1990 shown in Figure B1 reflect this tariffization process in which tariffs were raised to compensate for the protective effect of the import bans that were abolished in March 1990 (Carvalho 1992).

[31] Email correspondence with Honório Kume, March 12, 2008.

B3. Cross-Sectional Wage Regressions

In order to calculate the regional wage change for each microregion, I estimate standard wage regressions separately in 1991 and 2000. Wages are calculated as an individual's monthly earnings / 4.33 divided by weekly hours in their main job (results using all jobs are nearly identical). I regress the log wage on age, age-squared, a female indicator, an inner-city indicator, four race indicators, a marital status indicator, fixed effects for each year of education from 0 to 17+, 21 industry fixed effects, and 494 microregion fixed effects. By running these regressions separately by survey year, I control for changes in regional demographic characteristics and for changes in the national returns to those characteristics. The results of these regressions are reported in Table B2. All terms are highly statistically significant and of the expected sign. The regional fixed effects are normalized relative to the national average log wage, and their standard errors are calculated using the process described in Haisken-DeNew and Schmidt (1997). The regional wage change is then calculated as the difference in these normalized regional fixed effects between 1991 and 2000:

$$d \ln w_r = (\ln w_r^2 000 - \ln w_r^1 991) - (\overline{\ln w_r^{2000}} - \overline{\ln w_r^{1991}}) \tag{B1}$$

The regional wage changes are shown in Figure 2. As mentioned in the main text, some sparsely populated regions have very large measured regional wage changes. To demonstrate that these results are driven by the data and not some artifact of the wage regressions, Figure B3 shows a similar map plotting regional wage changes that were generated without any demographic or industry controls. They represent the change in the mean log wage in each region. The amount of variation across regions is similar, and these unconditional regional wage changes are highly correlated with the conditional versions used in the analysis, with a correlation coefficient of 0.93.

B4. Region-Level Tariff Change Elements

The fixed factor share of input costs is measured as one minus the labor share of value added in national accounts data. The resulting estimates of θ_i are plotted in Figure B4. The distribution of laborers across industries in each region (λ_{ri}) are calculated in the 1991 Census using the sample of employed individuals, not enrolled in school, aged 18-55. The tariff changes in 5 are calculated from the nominal tariff data in Kume, Piani and de Souza (2003) and are shown in Figure B5.

Figure B1. Nominal Tariff Timeline

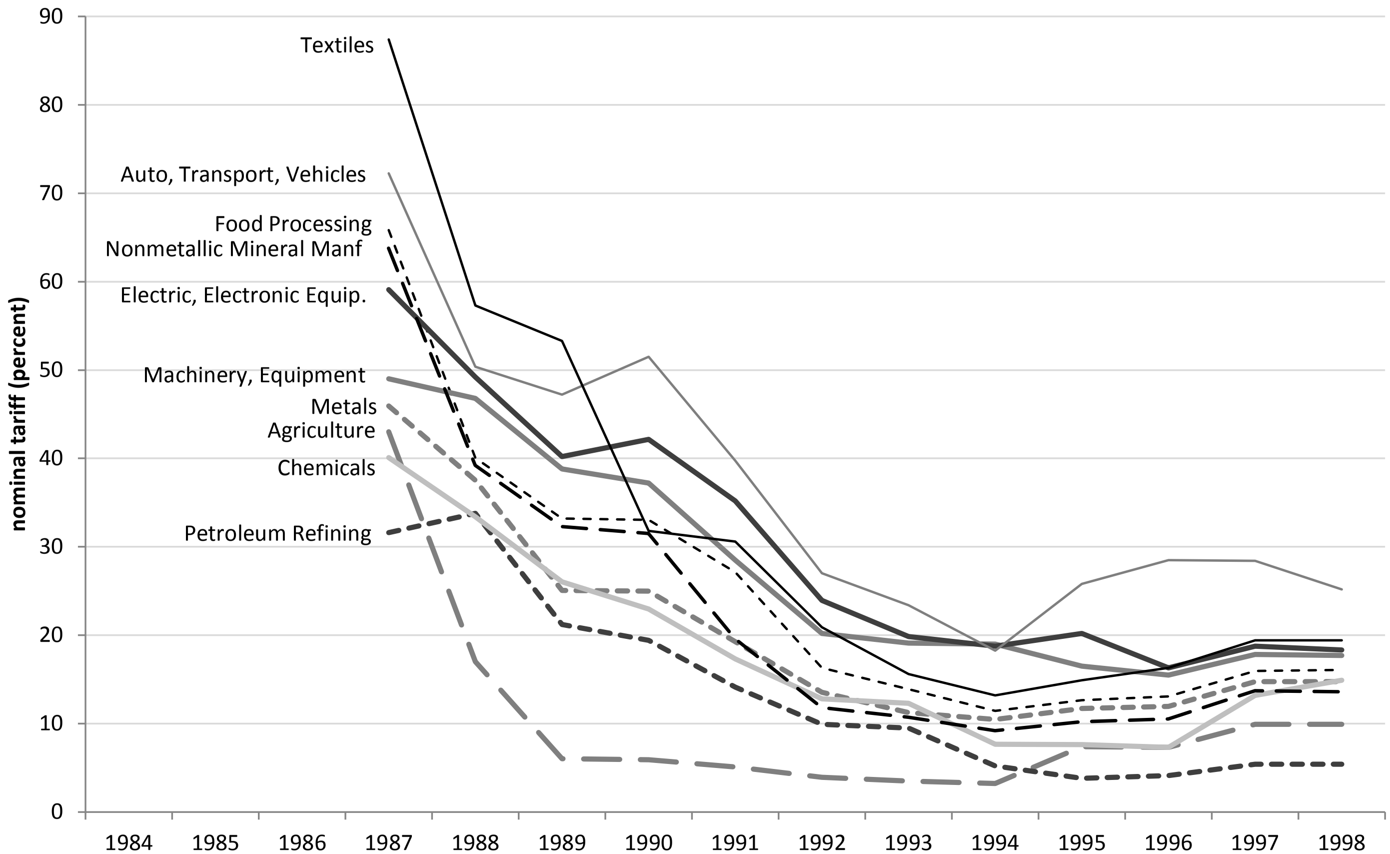


Figure B2. Effective Rate of Protection Timeline

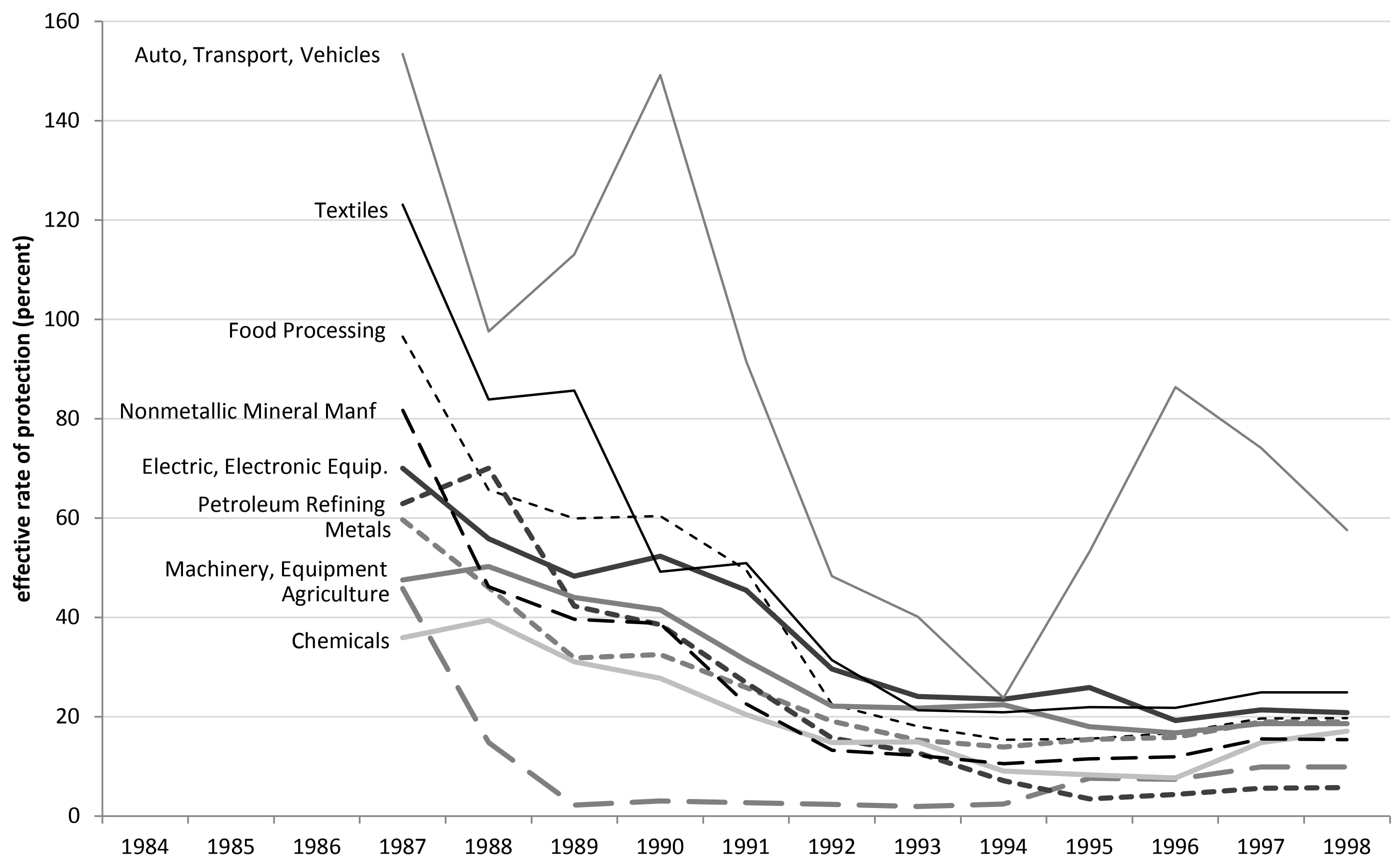
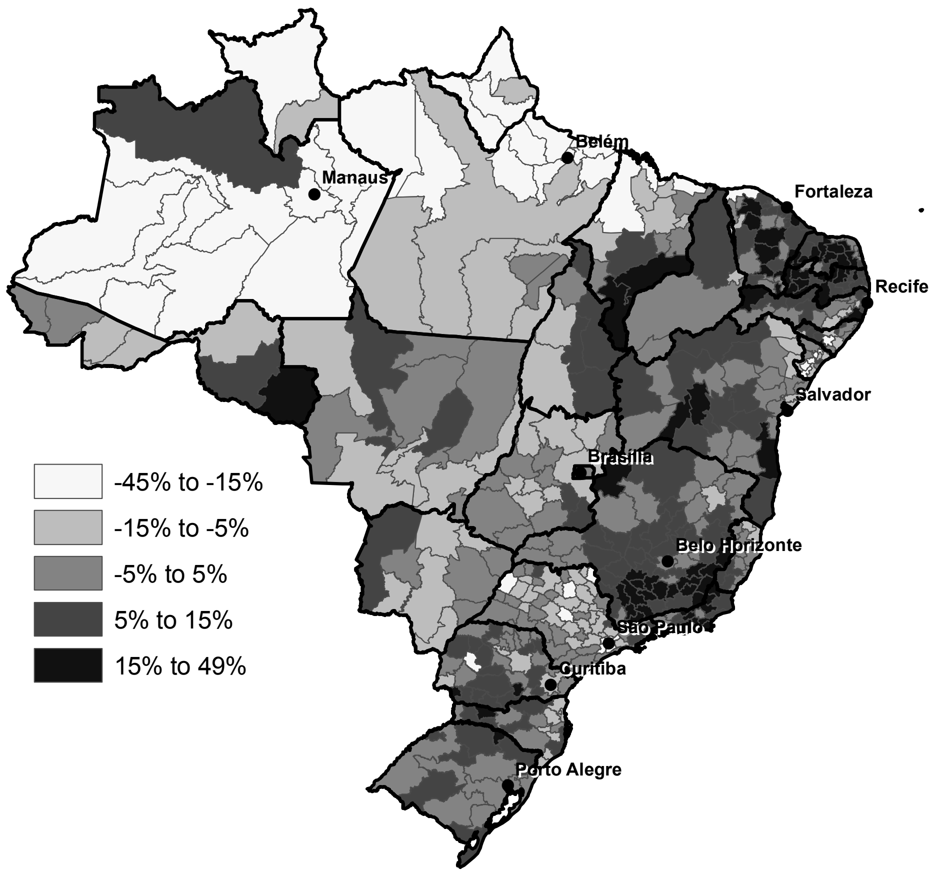
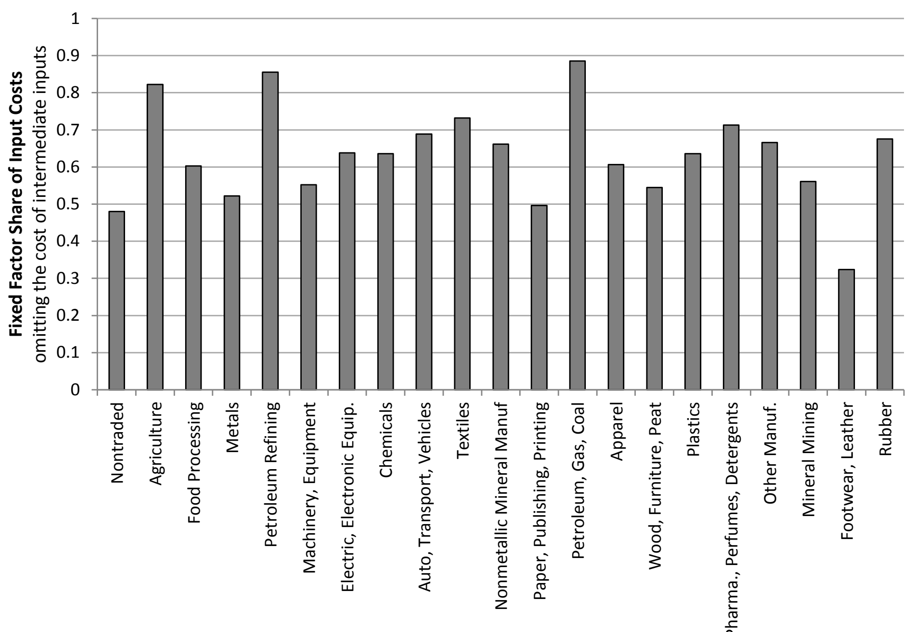


Figure B3. Unconditional Regional Wage Changes



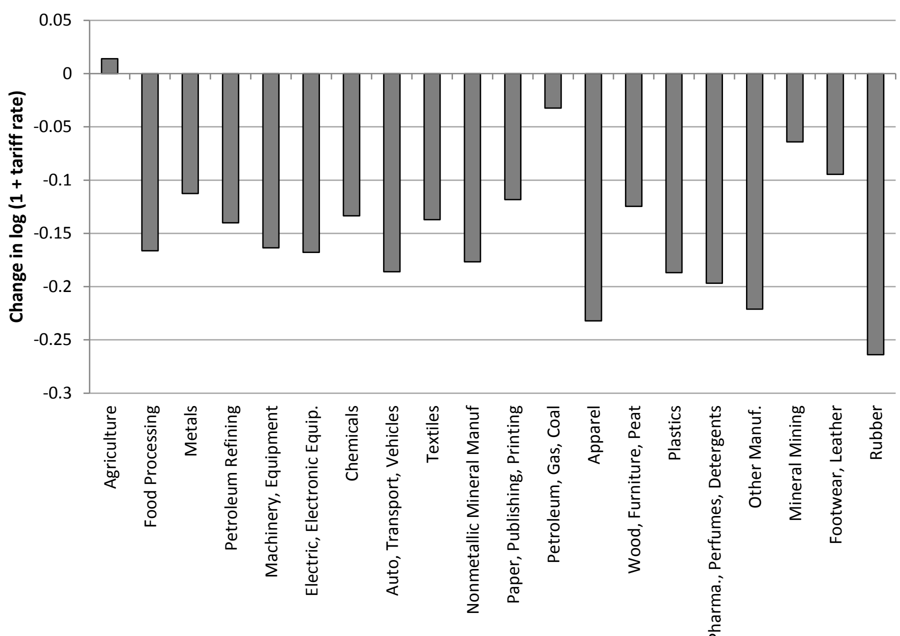
Proportional wage change by microregion - normalized change in microregion fixed effects without demographic controls

Figure B4. Fixed Factor Share of Input Costs



Source: IBGE National Accounts data 1990
Sorted by industry value added in 1990 (largest to smallest)

Figure B5. Tariff Changes



Source: Author's calculations based on Kume et al. (2003)
Sorted by industry value added in 1990 (largest to smallest)

Table B1—Industry Aggregation and Concordance

Final Industry	Sector Name	Nível 50	Nível 80	1991 Census (*atividade*)	2000 Census (*CNAE-Dom*)
1	Agriculture	1	101-199	011-037, 041, 042, 581	1101-1118, 1201-1209, 1300, 1401, 1402, 2001, 2002, 5001, 5002
2	Mineral Mining (except combustibles)	2	201-202	050, 053-059	12000, 13001, 13002, 14001-14004
3	Petroleum and Gas Extraction and Coal Mining	3	301-302	051-052	10000, 11000
4	Nonmetallic Mineral Goods Manufacturing	4	401	100	26010, 26091, 26092
5	Iron and Steel, Nonferrous, and Other Metal Production and Processing	5-7	501-701	110	27001-27003, 28001, 28002
8	Machinery, Equipment, Commercial Installation Manufacturing, and Tractor Manufacturing	8	801-802	120	29001
10	Electrical, Electronic, and Communication Equipment and Components Manufacturing	10-11	1001-1101	130	29002, 30000, 31001, 31002, 32000, 33003
12	Automobile, Transportation, and Vehicle Parts Manufacturing	12-13	1201-1301	140	34001-34003, 35010, 35020, 35030, 35090
14	Wood Products, Furniture Manufacturing, and Peat Production	14	1401	150, 151, 160	20000, 36010
15	Paper Manufacturing, Publishing, and Printing	15	1501	170, 290	21001, 21002, 22000
16	Rubber Product Manufacturing	16	1601	180	25010
17	Chemical Product Manufacturing	17,19	1701-1702, 1901-1903	200	23010, 23030, 23400, 24010, 24090
18	Petroleum Refining and Petrochemical Manufacturing	18	1801-1806	201, 202, 352, 477	23020
20	Pharmaceutical Products, Perfumes and Detergents Manufacturing	20	2001	210, 220	24020, 24030
21	Plastics Products Manufacturing	21	2101	230	25020
22	Textiles Manufacturing	22	2201-2205	240, 241	17001, 17002
23	Apparel and Apparel Accessories Manufacturing	23	2301	250,532	18001, 18002
24	Footwear and Leather and Hide Products Manufacturing	24	2401	190, 251	19011, 19012, 19020
25	Food Processing (Coffee, Plant Products, Meat, Dairy, Sugar, Oils, Beverages, and Other)	25-31	2501-3102	260, 261, 270, 280	15010, 15021, 15022, 15030, 15041-15043, 15050, 16000
32	Miscellaneous Other Products Manufacturing	32	3201	300	33001, 33002, 33004, 33005, 36090, 37000
99	Nontraded Goods and Services	33-43	3301-4301	340, 351-354, 410-424, 451-453, 461, 462, 464, 471-476, 481, 482, 511, 512, 521-525, 531, 533, 541-545, 551, 552, 571-578, 582-589, 610-619, 621-624, 631, 632, 711-717, 721-727, 901, 902	1500, 40010, 40020, 41000, 45001-45005, 50010, 50020, 50030, 50040, 50050, 53010, 53020, 53030, 53041, 53042, 53050, 53061-53068, 53070, 53080, 53090, 53101, 53102, 53111-53113, 55010, 55020, 55030. 60010. 60020. 60031. 60032. 60040. 60091. 60092, 61000, 62000, 63010, 63021, 63022, 63030, 64010, 64020, 65000, 66000, 67010, 67020, 70001, 70002, 71010, 71020, 71030, 72010, 72020, 73000, 74011, 74012, 74021, 74022, 74030, 74040, 74050, 74060. 74090. 75011-75017. 75020. 80011. 80012. 80090, 85011-85013, 85020, 85030, 90000, 91010, 91020, 91091, 91092, 92011-92015, 92020, 92030, 92040, 93010, 93020, 93030, 93091, 93092, 95000, 99000

Table B2—Cross-Sectional Wage Regressions - 1991 and 2000 Census

dependent variable: log wage = ln((monthly earnings / 4.33) / weekly hours) at main job

Year	1991	2000
Age	0.057	0.061
	(0.000)**	(0.000)**
Age^2 / 1000	-0.575	-0.601
	(0.004)**	(0.004)**
Female	-0.392	-0.335
	(0.001)**	(0.001)**
Inner City	0.116	0.101
	(0.001)**	(0.001)**
Race		
Brown (parda)	-0.136	-0.131
	(0.001)**	(0.001)**
Black	-0.200	-0.173
	(0.002)**	(0.001)**
Asian	0.154	0.122
	(0.006)**	(0.006)**
Indigenous	-0.176	-0.112
	(0.010)**	(0.006)**
Married	0.186	0.163
	(0.001)**	(0.001)**
Fixed Effects		
Years of Education (18)	X	X
Industry (21)	X	X
Microregion (494)	X	X
Observations	4,721,996	5,135,618
R-squared	0.518	0.500

Robust standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%
Omitted category: unmarried white male with zero years of education, outside inner city, working in agriculture