Online Appendix
The Impact of Trade Liberalization in the Presence of Political Distortions

By Sebastian Jävervall and Roza Khoban

A Measuring Political Connections

A.1 Surname Matching between Firm Directors and Politicians

Our primary measure of political connections is the surname matching between firm directors and politicians within an assembly constituency. To match surnames, we employ a fuzzy matching algorithm that accounts for the differential spellings of names. The algorithm, `masala_merge`, is based on the Levenshtein edit distance algorithm and is specialized for Hindi names to give smaller penalties for very common alternate spellings of names in Hindi.[1] Our main objective when using the algorithm is to match surnames that are either exactly the same or different only in their spelling.

We begin by processing and cleaning politician and director surnames by removing first names, titles, and abbreviations. We then depart from our data on the boards of directors for every firm-year pair and fuzzy match directors' surnames with the political data containing the surnames of the winner and runner-up candidate in every constituency-election year pair. The procedure yields a set of matches with associated *Levenshtein distances* indicating the match quality. The greater the distance is, the greater the number of additions, deletions, or insertions required to go from one surname to another. A distance of 0 indicates an exact surname match. The algorithm has a baseline cutoff that discards matches with high Levenshtein distances, indicating that two names are very unlikely to refer to the same surname. We use a conservative cutoff of the Levenshtein distance equal to 1 as our baseline measure.

In the second step, we manually check the set of probable matches. This is necessitated by the fact that surname matches with a low, but nonzero, Levenshtein distance may or may not be the same surname. For example, we want to keep cases where surnames such as "Reddy" and "Redddy" are matched as they are the same surname but with different spellings, but we want to discard matches such as "Raj" and "Rao", which have low Levenshtein distances but are in fact distinct surnames. To guide this decision, we process the set of probable surname matches through OpenAI's GPT-4. As a robustness check, we investigate the sensitivity of our results by using only exact surname matches

[1] `masala_merge` was developed by Paul Novosad. See https://github.com/devdatalab/masala-merge.

(Levenshtein distance equal to 0).

A.2 Feature Selection Measure of Political Connections

We cross-validate our main social proximity measure by additionally defining firms as politically connected using a data-driven χ^2 feature selection method. As discussed in the paper, the χ^2 method draws on the finding that boards of Indian firms are overwhelmingly homogeneous in terms of social status and religion (Bhagavatula et al., 2023). This suggests that if two surnames appear on the same board, the respective individuals are more likely to belong to the same social community. Using data on board members and exploiting the probability of two surnames appearing on the same board, we categorize whether two surnames are dependent, i.e., connected to the same community, rather than randomly distributed across boards.

To categorize whether two surnames are dependent rather than randomly distributed across boards, we calculate Pearson's χ^2 statistic for each surname pair (m, n) in district d:[2]

$$\chi_d^2(m,n) = \sum_{e_m \in \{0,1\}} \sum_{e_n \in \{0,1\}} \frac{(N^d_{e_m e_n} - E^d_{e_m e_n})^2}{E^d_{e_{md} e_{nd}}}$$

where e_m (e_n) is equal to 1 if a board in a district contains surname m (n) and 0 otherwise. $N^d_{e_m e_n}$ counts the observed number of occurrences of e_m and e_n in district d, while $E^d_{e_m e_n}$ calculates the expected number of such occurrences assuming independence of the two surnames. We again employ the fuzzy matching algorithm specialized for Indian names, discussed in online Appendix Section A.1, to account for differences in spelling.[3]

If the surname counts are drawn from (possibly different) multinomial distributions, χ^2 is a test statistic for the null hypothesis that the two surnames are independent. Thus, if we can reject the null hypothesis, the surname pair is categorized as dependent. We use the χ^2 calculations, with a cutoff p-value of 1%, to additionally define a firm as connected to a politician if any of the firm's board members' surnames and the politician's surname are defined as dependent by the feature selection method.[4]

[2]The χ^2 statistic can be calculated at both a lower (e.g., city or postal code) and higher (e.g., state) geographical level. The choice of district reflects a trade-off between bias in surname matching and precision.

[3]As the high number of possible matches makes it difficult to manually verify the output from the fuzzy matching algorithm, here, we use 0.8 as the cutoff to define matches.

[4]We use the common rule of thumb for χ^2 statistic validity and use only χ^2 calculations for surname pairs where there are at least 10 data points ($\sum_{e_m \in \{0,1\}} \sum_{e_n \in \{0,1\}} N^d_{e_m e_n} > 10$) and where each count of the observed number of occurrences has at least five occurrences ($N^d_{e_m e_n} > 5$).

B Additional Figures

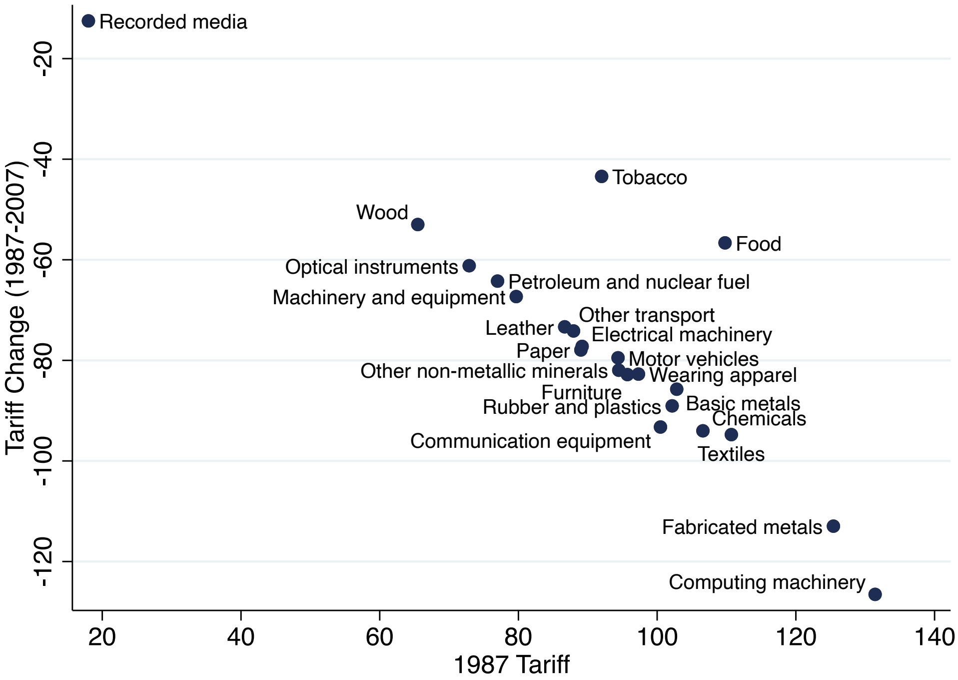


Figure B1: Prereform Tariffs and Tariff Changes: Sector Level

Notes: This figure shows a scatterplot of prereform tariff rates (measured in 1987) and the tariff change between 1987 and 2007 across sectors. Each marker represents a 2-digit NIC sector. “Food” is manufacture of food products and beverages. “Tobacco” is manufacture tobacco products. “Textiles” is manufacture of textiles. “Wearing apparel” is manufacture of wearing apparel; dressing and dyeing of fur. “Leather” is manufacture of tanning and dressing of leather; manufacture of luggage, handbags saddlery, harness, and footwear. “Wood” is manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plating materials. “Paper” is manufacture of paper and paper products. “Recorded media” is publishing, printing and reproduction of recorded media. “Petroleum and nuclear fuel” is manufacture of coke, refined petroleum products, and nuclear fuel. “Chemicals” is manufacture of chemicals and chemical products. “Rubber and plastics” is manufacture of rubber and plastic products. “Other non-metallic minerals” is manufacture of other non-metallic mineral products. “Basic metals” is manufacture of basic metals. “Fabricated metals” is manufacture of fabricated metal products, except machinery and equipment. “Machinery and equipment” is manufacture of machinery and equipment n.e.c. “Computing machinery” is manufacture of office, accounting, and computing machinery. “Electrical machinery” is manufacture of electrical machinery and apparatus n.e.c. “Communication equipment” is manufacture of radio, television, and communication equipment and apparatus. “Optical instruments” is manufacture of medical, precision and optical instruments, watches and clocks. “Motor vehicles” is manufacture of motor vehicles, trailers and semi-trailers. “Other transport” is manufacture of other transport equipment. “Furniture” is manufacture of furniture; manufacturing n.e.c.

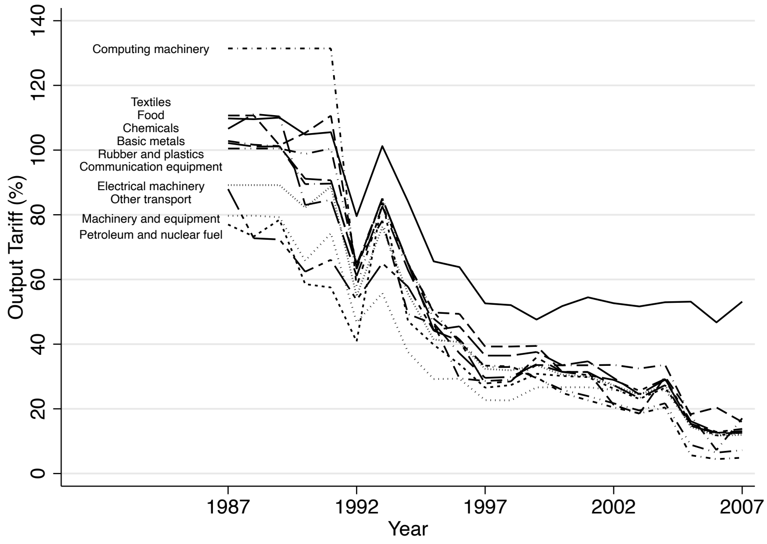


Figure B2: Output Tariffs, 1987–2007

Notes: This figure shows the evolution of the output tariff rates for the 11 largest 2-digit NIC sectors by 1990 value added. "Computing machinery" is manufacture of office, accounting, and computing machinery. "Textiles" is manufacture of textiles. "Food" is manufacture of food products and beverages. "Chemicals" is manufacture of chemicals and chemical products. "Basic metals" is manufacture of basic metals. "Rubber and plastics" is manufacture of rubber and plastic products. "Communication equipment" is manufacture of radio, television and communication equipment, and apparatus. "Electrical machinery" is manufacture of electrical machinery and apparatus n.e.c. "Other transport" is manufacture of other transport equipment. "Machinery and equipment" is manufacture of machinery and equipment n.e.c. "Petroleum and nuclear fuel" is manufacture of coke, refined petroleum products, and nuclear fuel.

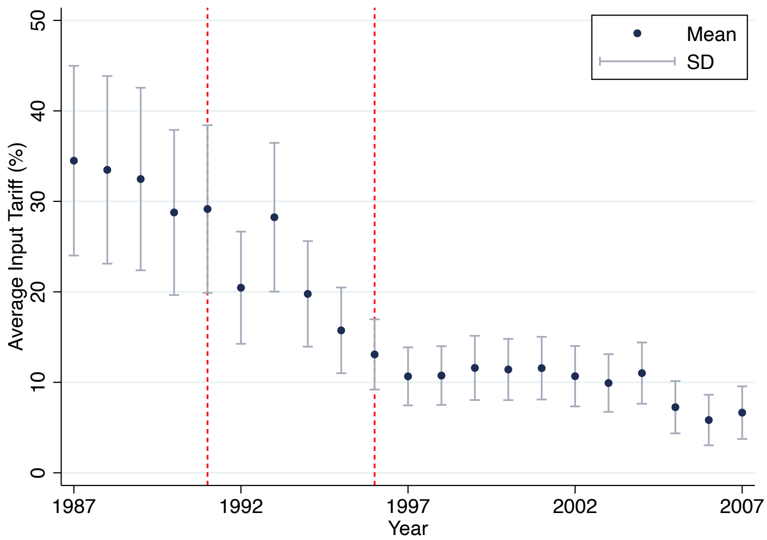
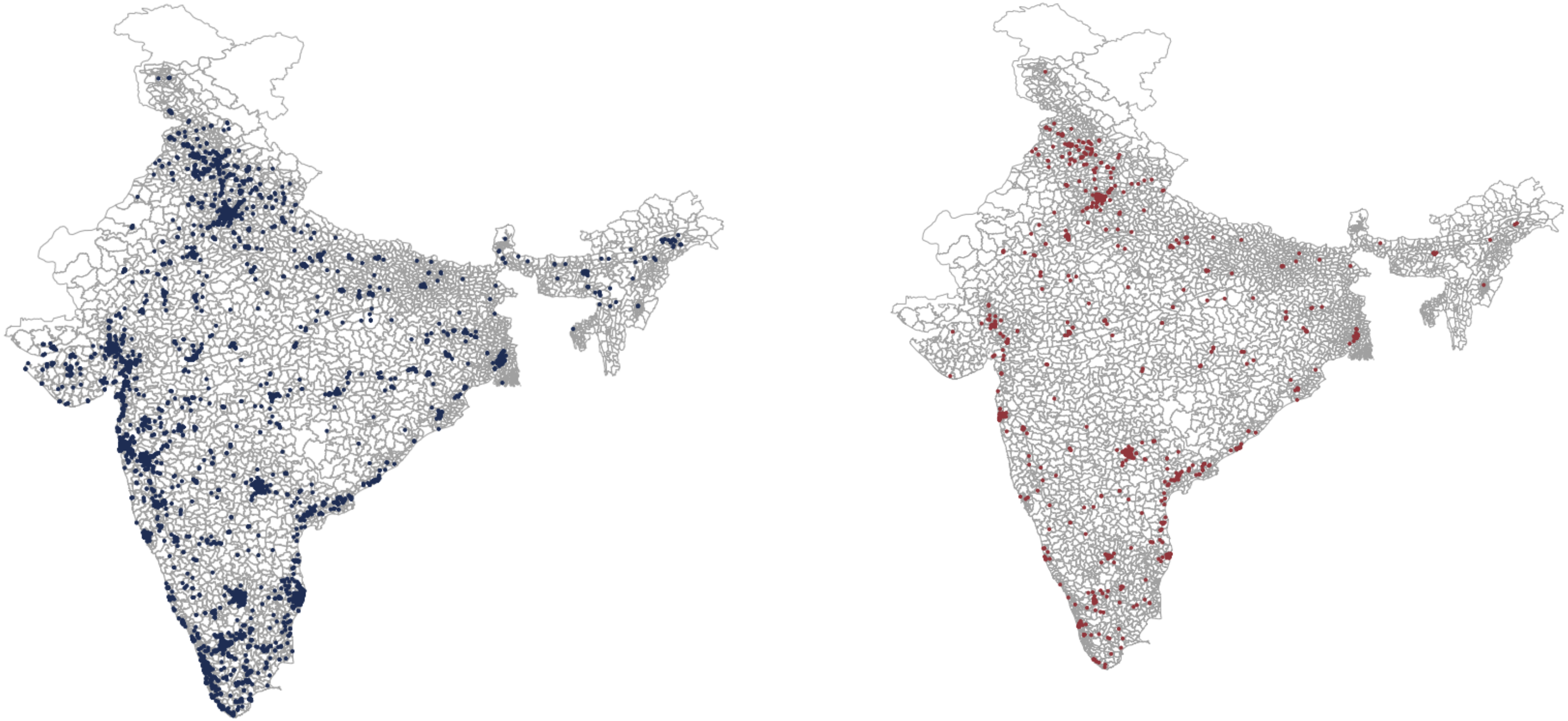


Figure B3: Average Input Tariff, 1987–2007

Notes: This figure shows the evolution of the average input tariff rate and the standard deviation of input tariffs over the Indian trade liberalization period. The first vertical dotted line indicates 1991, the year the liberalization started. The period between 1991 and 1996 denotes the most intensive period of tariff reductions.



(a) All Firms

(b) Connected Firms

Figure B4: Firm Location and Constituency Boundaries

Notes: This figure shows maps of Indian assembly constituencies in India's predelimitation period. The borders represent assembly constituencies and markers represent the location of a firm's headquarters. Panel (a) includes the full sample of firms (24,396 firms), and panel (b) includes the sample of firms connected to either winner or runner-up politicians (2,236 firms).
Source: Constituency boundaries: ML InfoMap (2008).

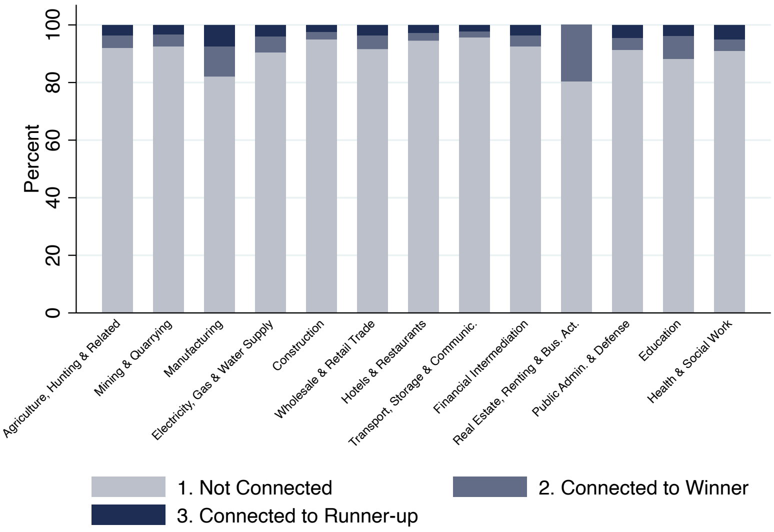


Figure B5: Distribution of Connections by Sector

Notes: This figure shows the distribution of firms' connection status by sectors (the 1-digit NIC level).

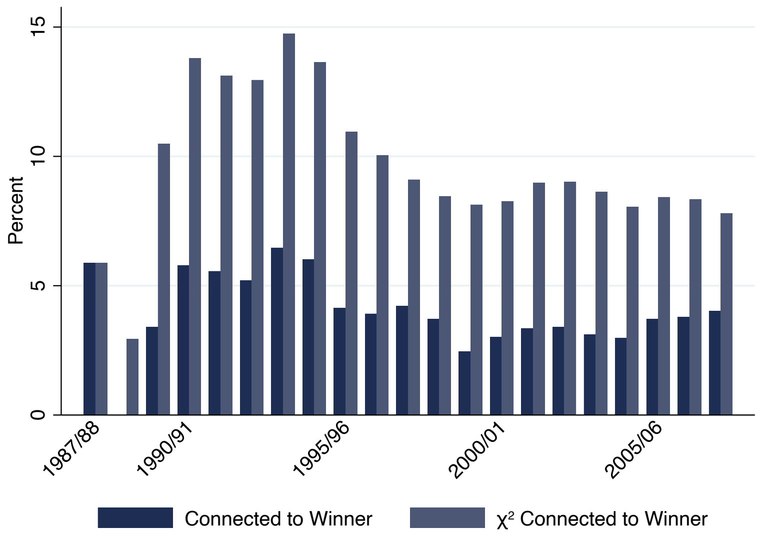
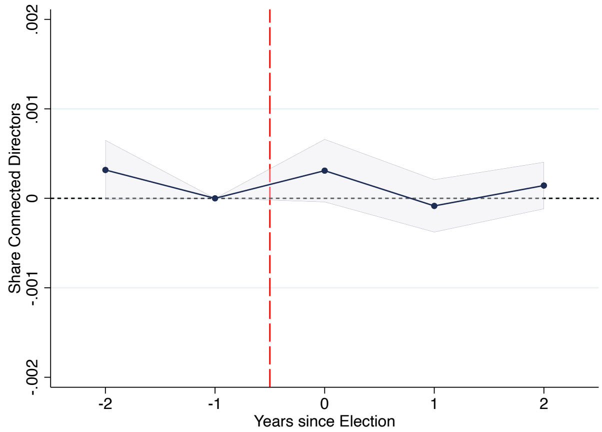
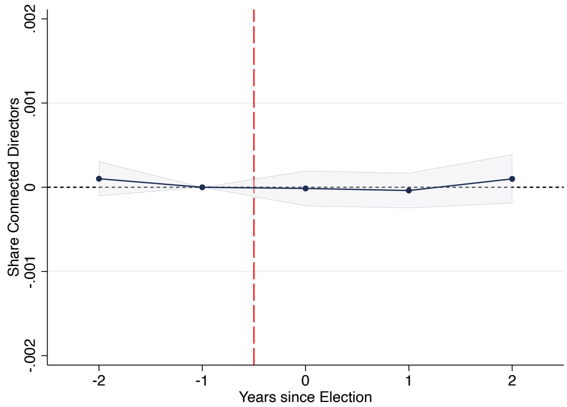


Figure B6: Distribution of Firms Connected to MLAs by Year

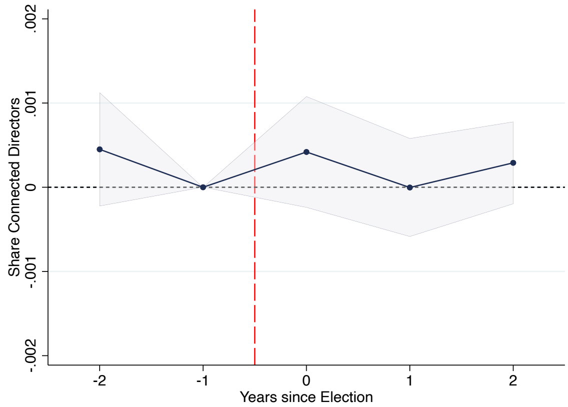
Notes: This figure shows the distribution of firm connections to winning politicians (MLAs) by year. *Connected to Winner* reports the percent of firms connected to the winning politician. χ^2 *Connected to Winner* reports the percent of firms connected to the winning politician according to the χ^2 definition.



(a) Full Period



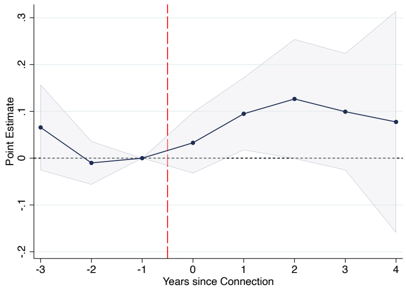
(b) 1987–1996



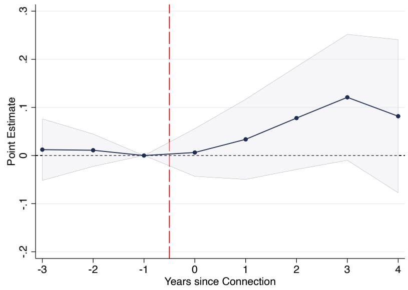
(c) 1997–2007

Figure B7: Share of Directors in Constituency Connected to the MLA

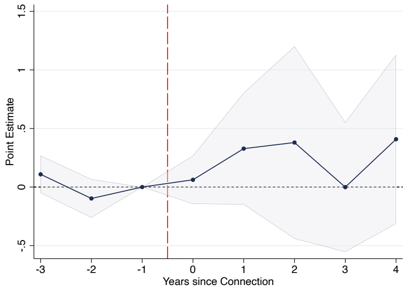
Notes: This figure shows the share of directors who are connected to the winning politician before and after the politician is elected. The specification is $ShareDirConnected_{pct} = \sum_{g=-2}^{2} \beta_g \mathbb{1}(t = E_i + g) + \alpha_c + \gamma_t + \xi_{pct}$, where $ShareDirConnected_{pct}$ is the share of directors in constituency c that are connected to the politician who won the election at $g = 0$. Panel (a) shows the estimated share of firms becoming connected to the winning politician for the full sample period. Panels (b) and (c) restrict the analysis by time period. The year prior to the election is normalized to zero. 95% confidence intervals are used for inference. Standard errors are clustered by constituency.



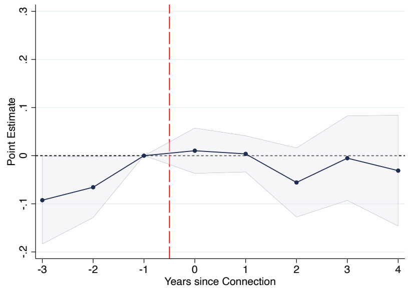
(a) Log Sales



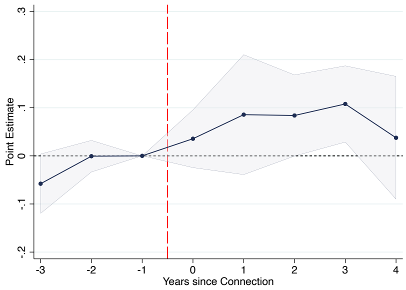
(b) Log Expenses



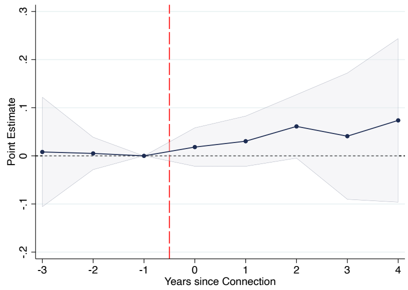
(c) Profit



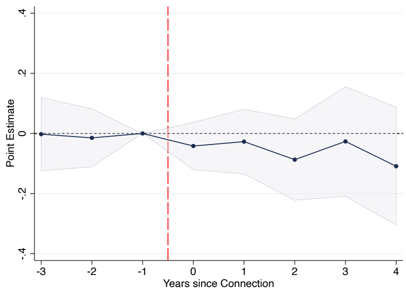
(d) TFP



(e) Log Capital



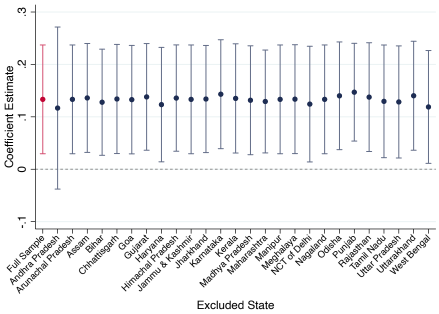
(f) Log Wage Bill



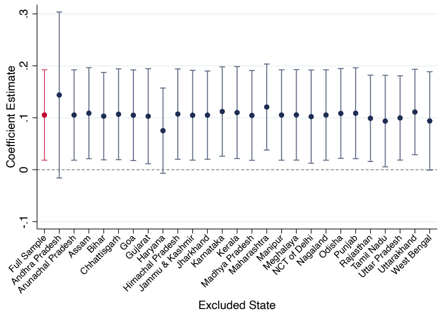
(g) Log Material

Figure B8: Dynamic Effects of Political Connections: Alternative Estimator

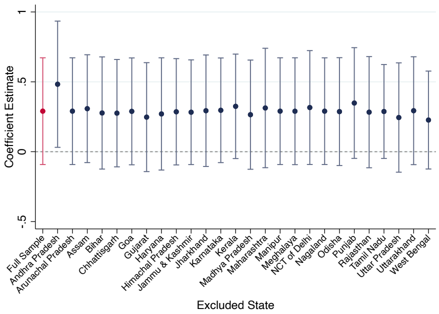
Notes: This figure illustrates the dynamic impact of political connections and examines pretrends by using the method proposed by de Chaisemartin and D'Haultfoeuille (2022). 95% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.



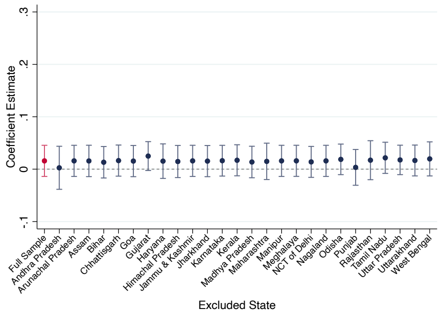
(a) Log Sales



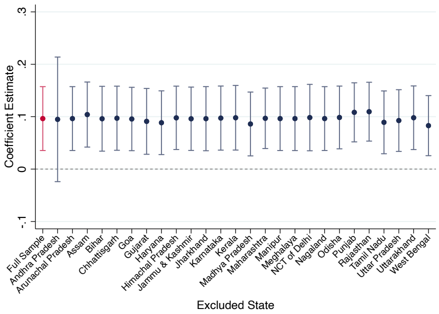
(b) Log Expenses



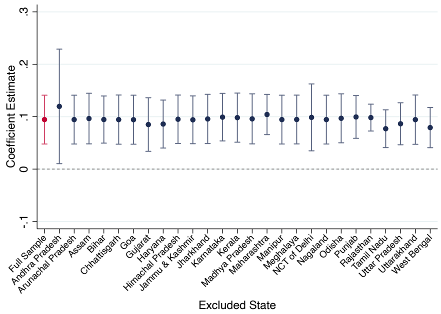
(c) Profit



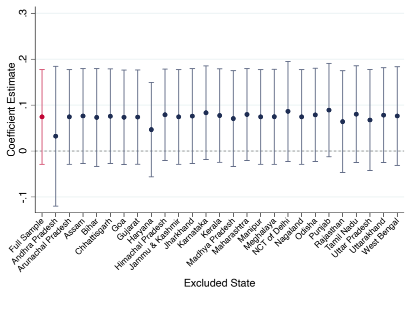
(d) TFP



(e) Log Capital



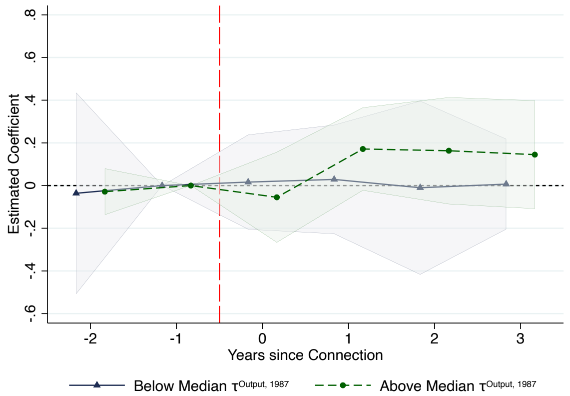
(f) Log Wage Bill



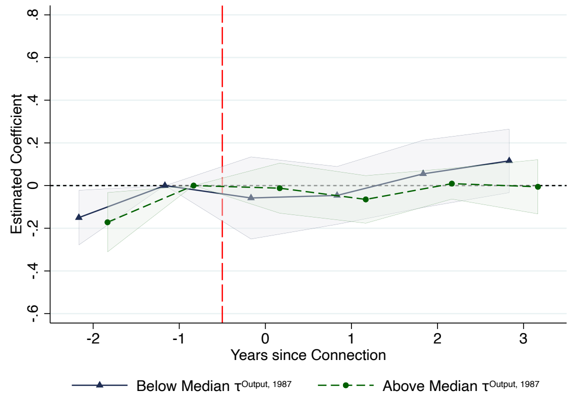
(g) Log Material

Figure B9: Effects of Political Connections: Excluding One State at a Time

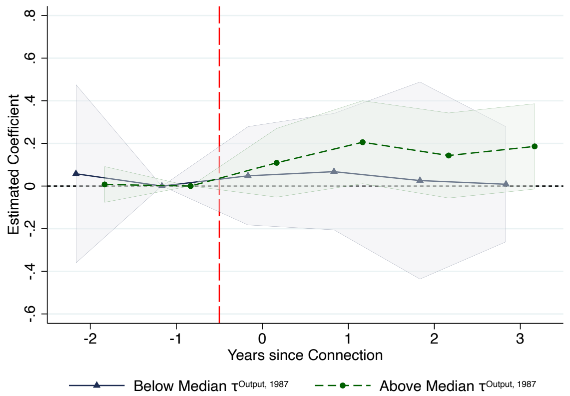
Notes: This figure shows the results from estimating equation (1) when excluding one state at a time from estimation. The figure shows coefficient estimates of the effect of $Connected_{ict}$ with associated 95% confidence intervals. Standard errors are two-way clustered by industry and constituency.



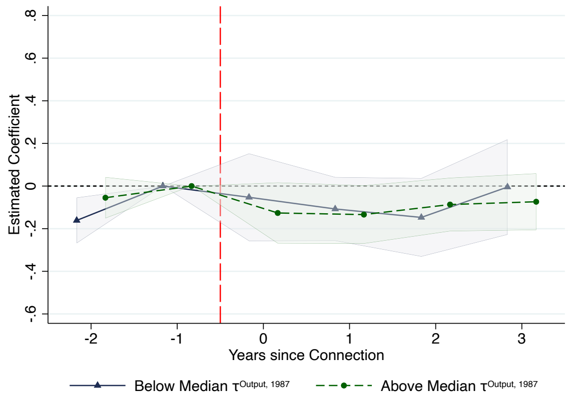
(a) Log Sales (1987–1996)



(b) Log Sales (1997–2007)



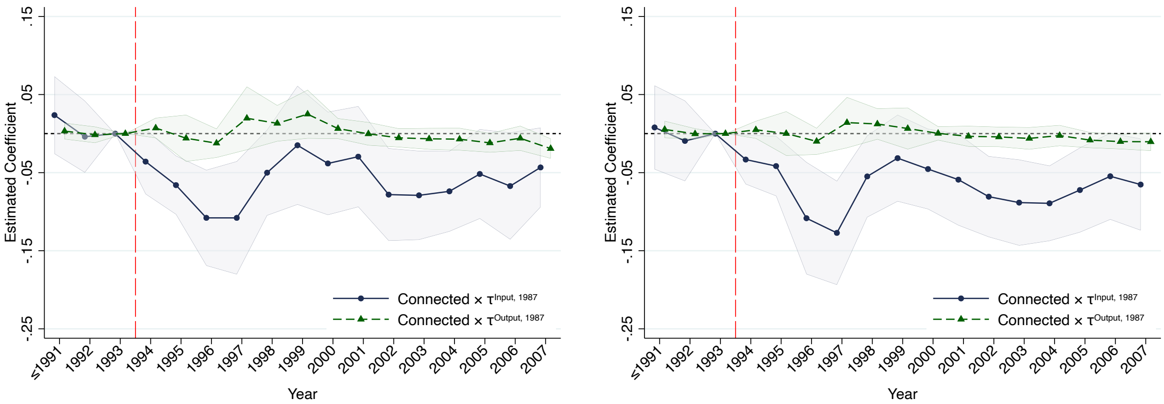
(c) Log Material (1987–1996)



(d) Log Material (1997–2007)

Figure B10: Dynamic Effects of Political Connections by Period and Prereform Trade Protection: Output Tariffs

Notes: This figure illustrates the dynamic impact of political connections by time period and degree of prereform trade protection by plotting estimates ($\hat{\pi}_j$) from equation (2). The sample is split into firms in industries with output tariffs below the median output tariff level in 1987 and firms in industries with output tariffs above the median output tariff level in 1987. Panels (a) and (b) show the effect on (log) firm sales, where panel (a) reports estimates for each of the two groups separately in 1987–1996, and panel (b) reports estimates for each of the two groups separately in 1997–2007. Panels (c) and (d) report the equivalent estimates when using (log) materials as the outcome variable. The year prior to the connection event is normalized to 0. 95% confidence intervals are used for inference, and standard errors are two-way clustered by industry and constituency.

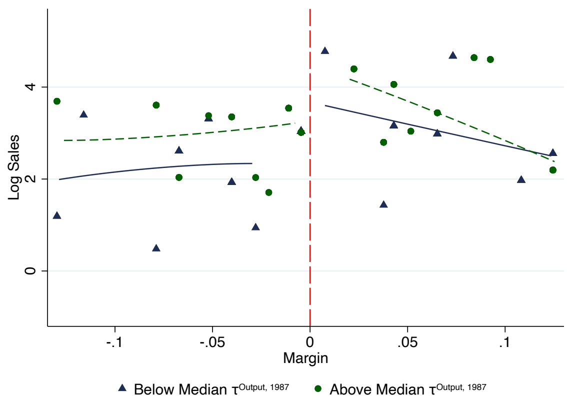


(a) Log Sales

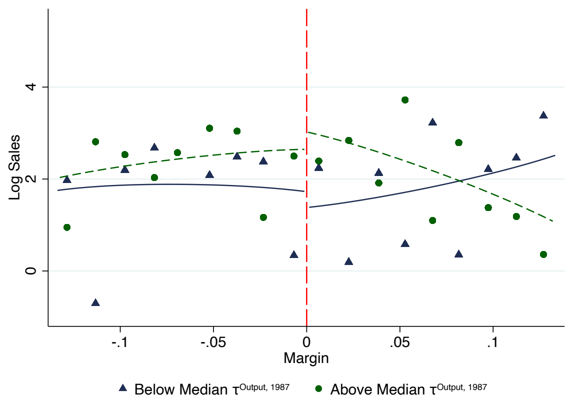
(b) Log Material

Figure B11: Triple-Differences Design: Pretrends and Dynamic Impacts

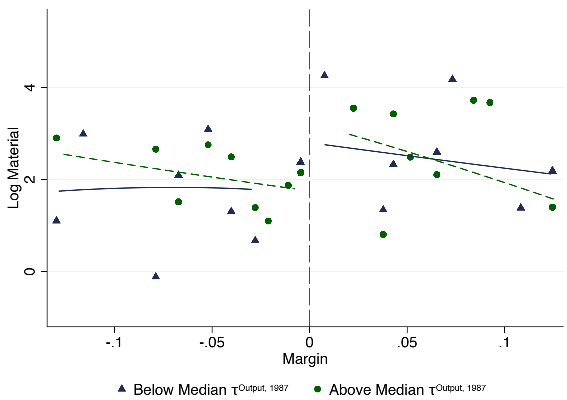
Notes: This figure shows yearly impacts of tariff reductions on returns to political connections by plotting the $\hat{\pi}_t$ coefficients from the regression $y_{ikct} = \beta Connected_{ict} + \boldsymbol{\tau}'_{k87}\vartheta Connected_{ict} + \sum_{t=1987,t\neq1993}^{2007}\delta_t\mathbb{1}(year = t) \times Connected_{ict} + \sum_{t=1987,t\neq1993}^{2007}\boldsymbol{\tau}'_{k87}\theta_t\mathbb{1}(year = t) + \sum_{t=1987,t\neq1993}^{2007}\pi_t\mathbb{1}(year = t) \times \boldsymbol{\tau}'_{k87}Connected_{ict} + \alpha_i + \boldsymbol{\gamma}\mathbf{X}_{ikct} + \epsilon_{ikct}$. $\boldsymbol{\tau}_{k87} = (\tau_{87}^{Output}, \tau_{87}^{Input})$ denotes a vector of demeaned baseline (1987) output and input tariff rates for industry k. The dependent variable in panel (a) is (log) sales, and the dependent variable in panel (b) is (log) material. 1993 is normalized to 0. 90% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.



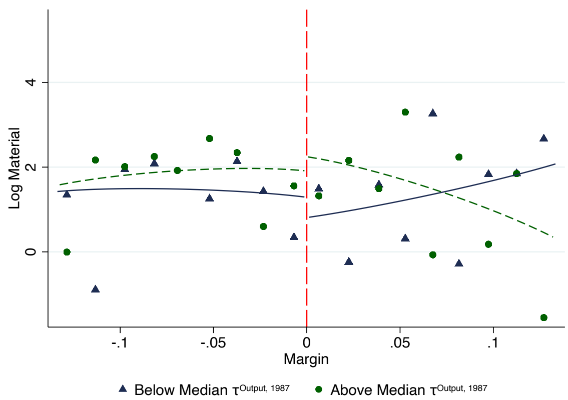
(a) Log Sales (1987–1996)



(b) Log Sales (1997–2007)



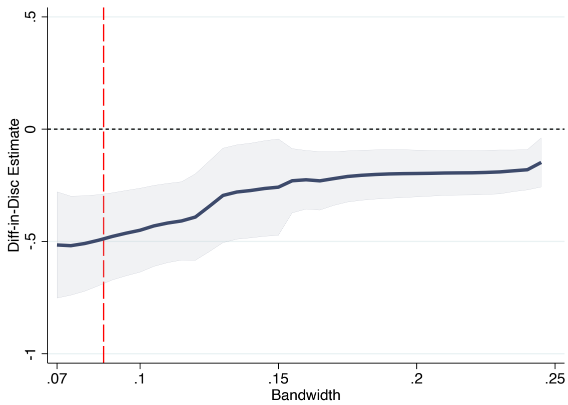
(c) Log Material (1987–1996)



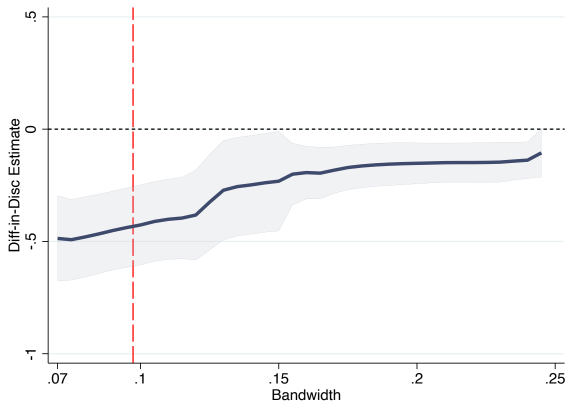
(d) Log Material (1997–2007)

Figure B12: Illustration of the Difference-in-Discontinuities Design: By Output Tariff

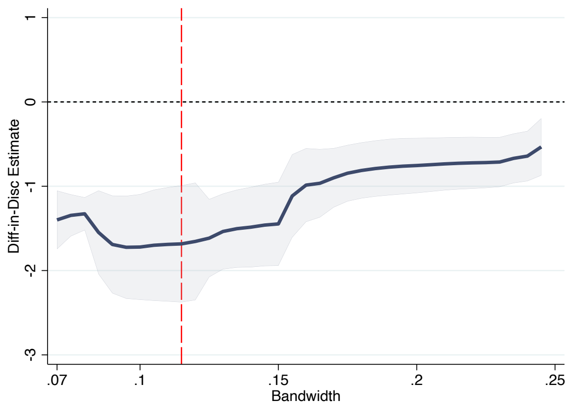
Notes: This figure illustrates how the effect of political connections changes in the pre/during and post-reform periods for firms in industries with above- and below-median output tariffs in 1987. Panels (a) and (b) show the effect on firms' (log) sales, where panel (a) estimates discontinuities in 1987–1996, and panel (b) estimates discontinuities in 1997–2007. Panels (c) and (d) illustrate the effects on firms' (log) material expenses. The main specification (equation (6)) uses the full, continuous variation in tariffs, while this figure splits the full sample of firms into groups with baseline tariffs above and below the median. Therefore, the figure should be read only as an approximation to the identifying variation in the regression results reported in subsequent tables. The curves are local linear regressions with an Epanechnikov kernel and bandwidth of 8%. The estimation window is restricted to 13.5% on each side of the threshold.



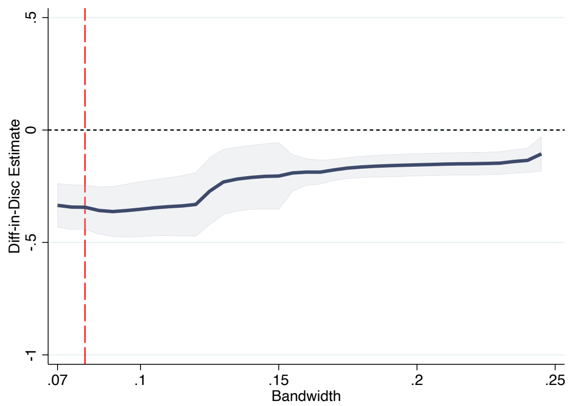
(a) Log Sales



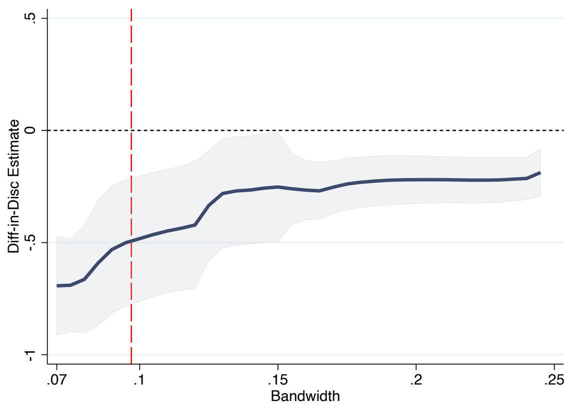
(b) Log Expenses



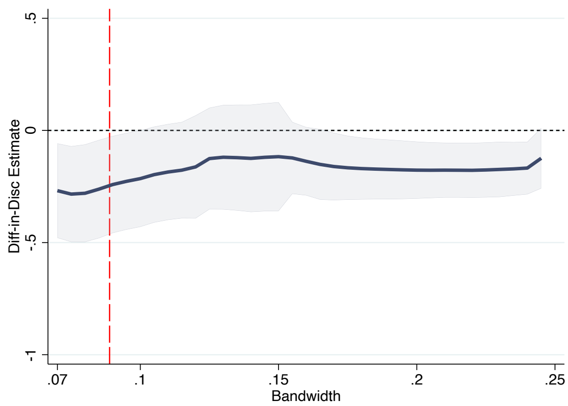
(c) Profit



(d) Log Capital



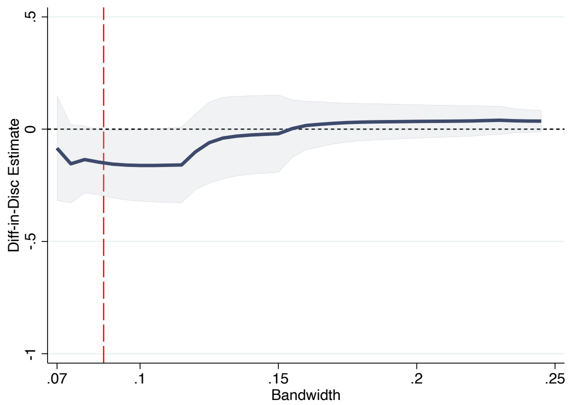
(e) Log Wage Bill



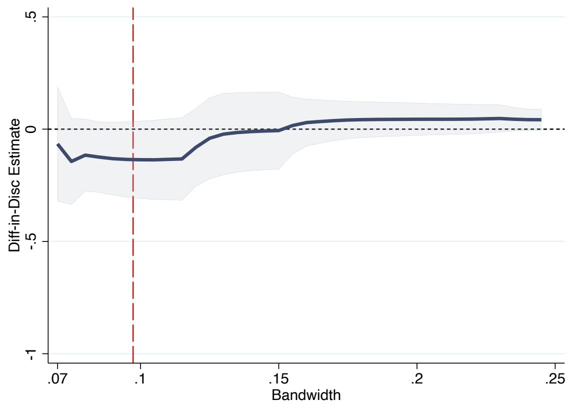
(f) Log Material

Figure B13: Bandwidth Sensitivity of Diff-in-Disc Estimates: $Connected_{ict} \times Post_t \times \tau_{k87}^{Input}$

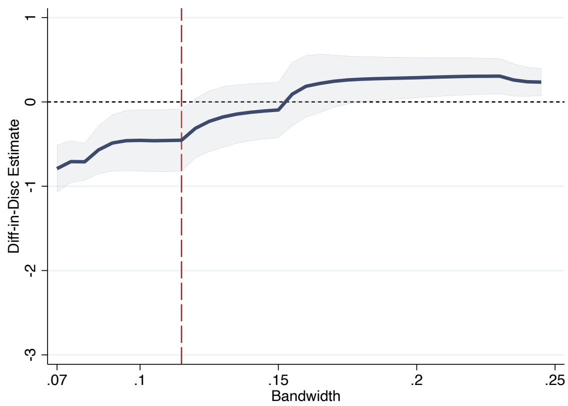
Notes: This figure shows estimated Diff-in-Disc coefficients for different bandwidths. The specification is described in Section 5. The vertical dotted line indicates the optimal bandwidth (Calonico, Cattaneo and Titiunik, 2014). 95% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.



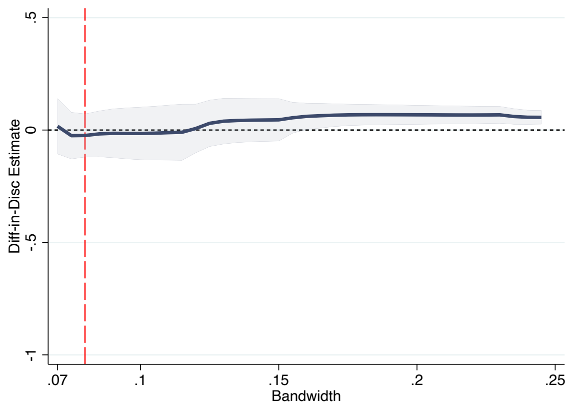
(a) Log Sales



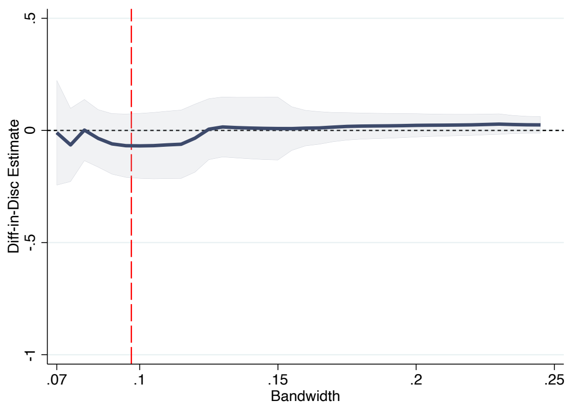
(b) Log Expenses



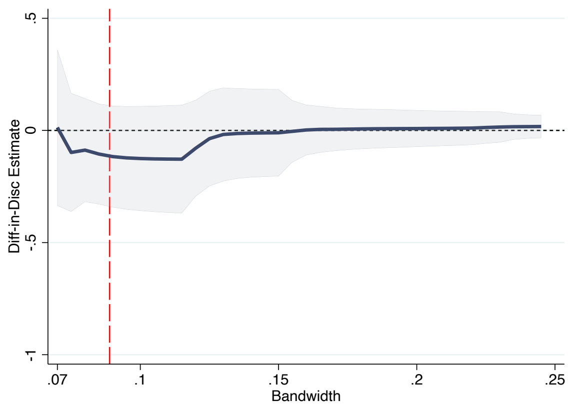
(c) Profit



(d) Log Capital



(e) Log Wage Bill



(f) Log Material

Figure B14: Bandwidth Sensitivity of Diff-in-Disc Estimates: $Connected_{ict} \times Post_t \times \tau_{k87}^{Output}$

Notes: This figure shows estimated Diff-in-Disc coefficients for different bandwidths. The specification is described in Section 5. The vertical dotted line indicates the optimal bandwidth (Calonico, Cattaneo and Titiunik, 2014). 95% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.

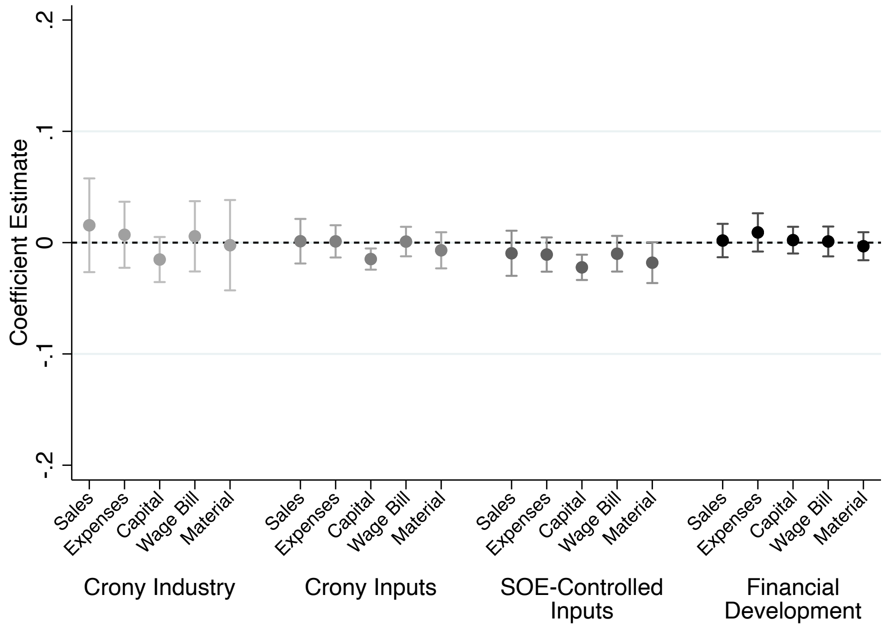


Figure B15: Heterogeneity in Output Tariff Effects

Notes: This figure shows estimated heterogeneities in the effects of tariff declines by industry and state characteristics. All dependent variables are in logs. Markers display effects of $Connected_{ict} \times Post_t \times \tau_{k87}^{Output} \times Z$, where Z is the factor of heterogeneity. 90% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.

C Additional Tables

Table C1: Summary Statistics: Elections and Politicians

	All Politicians			Connected Politicians			Close Elections Sample			Comparison (1)–(4)		Comparison (4)–(7)	
	Mean (1)	S.D. (2)	Obs. (3)	Mean (4)	S.D. (5)	Obs. (6)	Mean (7)	S.D. (8)	Obs. (9)	Diff. (10)	*p*-value (11)	Diff. (12)	*p*-value (13)
Winner	0.500	0.500	6,596	0.513	0.500	920	0.522	0.500	429	−0.013	0.459	−0.009	0.756
Margin	0.000	0.187	6,596	0.003	0.179	920	0.002	0.052	429	−0.003	0.628	0.001	0.865
Incumbent	0.223	0.416	6,596	0.237	0.425	920	0.238	0.426	429	−0.014	0.348	−0.001	0.974
Female Candidate	0.066	0.248	6,596	0.048	0.214	920	0.042	0.201	429	0.018	0.035	0.006	0.632
Aligned Candidate	0.467	0.499	6,596	0.472	0.499	920	0.473	0.500	429	−0.005	0.792	−0.001	0.960
INC Candidate	0.327	0.469	6,596	0.341	0.474	920	0.338	0.474	429	−0.014	0.398	0.003	0.905
BJP Candidate	0.188	0.391	6,596	0.202	0.402	920	0.179	0.384	429	−0.014	0.320	0.023	0.328
SC/ST Reserved	0.166	0.372	6,596	0.109	0.311	920	0.126	0.332	429	0.057	0.000	−0.017	0.356
Contestants	10.244	6.453	6,596	10.024	5.944	920	9.231	4.670	429	0.220	0.328	0.793	0.015
Turnout	0.630	0.110	6,596	0.625	0.119	920	0.647	0.121	429	0.005	0.235	−0.022	0.002

Notes: This table reports descriptive statistics of the electoral data from the SHRUG project (Jensenius and Verniers, 2017; Asher et al., 2021) for the period 1987–2007. *All Politicians* refers to the sample of winner and runner-up candidates in constituencies that can be spatially matched to Prowess firms. *Connected Politicians* refers to the sample of winner and runner-up candidates defined as connected to at least one firm by our surname matching methodology. *Close Elections Sample* refers to the sample of winner and runner-up candidates defined as connected to at least one firm by our surname matching methodology in close elections (as determined by the optimal bandwidth calculated using the method in Calonico, Cattaneo and Titiunik (2014)). Winner is a dummy indicating whether a candidate won the election in their constituency. Margin is the vote share difference between the two top-ranking candidates in each constituency. Incumbent is a dummy indicating whether a candidate won the preceding election. Female Candidate is a dummy indicating whether a candidate is female. Aligned Candidate is a dummy indicating whether a candidate represents a party that is part of the majority or coalition in the state government. INC Candidate is a dummy indicating whether a candidate represents the INC. BJP Candidate is a dummy indicating whether a candidate represents the BJP. SC/ST Reserved is a dummy indicating whether a constituency is reserved for SC/ST candidates. Contestants is the number of candidates standing for election in a constituency. Turnout is the ratio between actual votes cast and total number of eligible voters in a constituency.

Table C2: Summary Statistics: Firms

	All Firms			Connected Firms			Close Elections Sample		
	Mean (1)	S.D. (2)	Obs. (3)	Mean (4)	S.D. (5)	Obs. (6)	Mean (7)	S.D. (8)	Obs. (9)
Connected	0.036	0.187	118,990	0.582	0.493	7,381	0.577	0.494	2,482
Connected Winner or Runner-up	0.062	0.241	118,990	1.000	0.000	7,381	1.000	0.000	2,482
χ^2 Connected	0.086	0.281	118,990	0.606	0.489	7,381	0.629	0.483	2,482
χ^2 Connected Winner or Runner-up	0.135	0.342	118,990	1.000	0.000	7,381	1.000	0.000	2,482
τ_{k87}^{Output}	99.288	29.349	44,736	99.852	32.315	3,065	98.310	30.323	1,063
τ_{k87}^{Input}	37.700	7.093	44,491	36.338	7.832	3,047	36.917	7.742	1,057
$\Delta\tau_{k}^{Output}$	-47.227	22.064	44,736	-47.123	24.136	3,065	-46.648	22.609	1,063
$\Delta\tau_{k}^{Input}$	-17.647	4.125	44,491	-16.819	4.400	3,047	-17.104	4.348	1,057
Sales	52.008	564.469	65,914	59.263	363.078	4,933	82.104	565.226	1,635
Expenses	45.179	500.728	83,689	52.318	291.733	5,788	66.400	429.816	1,945
Profit	2.971	6.404	84,270	3.737	7.171	5,815	3.777	7.355	1,954
TFP	0.747	1.219	29,972	0.868	1.107	2,295	0.777	1.144	789
Capital	104.014	1444.477	89,004	123.005	1024.357	6,161	121.979	1190.734	2,055
Wage Bill	3.911	35.110	74,620	5.036	34.312	5,386	5.932	43.100	1,800
Material	36.163	481.940	52,360	33.961	153.240	3,910	44.985	217.697	1,323
Director Remuneration	0.087	0.361	30,073	0.111	0.404	2,528	0.096	0.373	842
Importer	0.265	0.441	118,990	0.324	0.468	7,381	0.336	0.472	2,482
Exporter	0.246	0.431	118,990	0.301	0.459	7,381	0.315	0.465	2,482
Import Share	0.028	0.088	83,599	0.029	0.088	5,784	0.028	0.088	1,945
Export Share	0.226	0.255	22,925	0.257	0.271	1,734	0.244	0.264	624
Age	19.983	17.945	118,990	20.201	18.636	7,381	22.380	21.744	2,482
No. of Directors	5.399	3.552	118,990	6.914	3.912	7,381	6.765	3.899	2,482
Private, Domestic	0.934	0.248	118,990	0.922	0.268	7,381	0.907	0.291	2,482
Private, Foreign	0.034	0.180	118,990	0.019	0.136	7,381	0.021	0.143	2,482
State Owned	0.032	0.177	118,990	0.059	0.235	7,381	0.073	0.259	2,482
Publicly Listed	0.343	0.475	118,990	0.447	0.497	7,381	0.458	0.498	2,482

Notes: This table reports statistics of the firm-level panel data from Prowess for the period 1987–2007. *All Firms* refers to the full sample of firms. *Connected Firms* refers to firms connected to either the winner or runner-up politician in their constituency. *Close Elections Sample* refers to the sample of firms connected to either the winner or the runner-up in their constituency in close elections (as determined by the optimal bandwidth calculated using the method in Calonico, Cattaneo and Titiunik (2014)). Connected is a dummy indicating whether a firm is connected to the winning politician in its constituency in a year. Connected Winner or Runner-up is a dummy indicating whether a firm is connected to the winning or runner-up politician. χ^2 Connected and χ^2 Connected Winner or Runner-up are the equivalents when using the χ^2 measure to define connections. τ_{k87}^{Output} and τ_{k87}^{Input} denote 1987 industry-level output and input tariff rates, measured in percent. $\Delta\tau_{k}^{Output}$ and $\Delta\tau_{k}^{Input}$ denote output and input tariff changes between 1991 and 1996. Sales sums all income generated from sales of goods and non-financial services. Total Expenses sums all revenue expenses incurred by the firm. Profit is the value left after meeting all expenses but before paying taxes. TFP is the revenue total factor productivity, estimated as suggested by Levinsohn and Petrin (2003). Capital is measured by total assets. Wage Bill is the remuneration in cash or in kind paid to employees. Material is the sum of expenses incurred on (i) raw material, (ii) stores, spares, and tools, (iii) finished goods, and (iv) packaging. Director remuneration is the sum of salaries, bonus/commissions, perquisites, retirement benefits, and other benefits paid to the directors of the firm. Importer is a dummy indicating whether a firm is importing, and Exporter is a dummy indicating whether a firm is exporting. Import share is the amount spent on imported inputs divided by total expenses. Export share is the earnings from exporting divided by total sales. Sales, expenses, profit, capital, wage bill, and material are measured in millions of USD. Age is the number of years since the firm was formed. No. of Directors is the size of the board of directors. Private, Domestic indicates whether a firm is categorized as owned by the Indian private sector, where the categorization reflects the structure of the ownership of the equity shares and the management control of the company. Private, Foreign indicates whether a firm is categorized as owned by foreigners, including foreign governments. State Owned indicates whether the firm is state-owned. Publicly Listed indicates whether the firm is publicly listed.

Table C3: Difference-in-Differences Estimates: Effect of Political Connections on Firm Size: Manufacturing Sample

	Sales	Expenses	Profit
	(1)	(2)	(3)
Panel A. Manufacturing Firms			
Connected	0.103	0.111**	0.036
	(0.065)	(0.051)	(0.179)
Panel B. Manufacturing Firms: χ^2 Measure			
χ^2 Connected	0.132**	0.108***	0.398*
	(0.053)	(0.039)	(0.227)
Observations	32,228	33,482	33,558
Year & Firm FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports the estimated effect of political connections on firm size outcomes, as specified by equation (1). Panel A reports results for manufacturing firms and panel B reports results for manufacturing firms using the alternative χ^2 definition of political connections. All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C4: Difference-in-Differences Estimates: Effect of Political Connections on Firm Production: Manufacturing Sample

	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)
Panel A. Manufacturing Firms				
Connected	0.016	0.069	0.077	0.030
	(0.015)	(0.045)	(0.049)	(0.048)
Panel B. Manufacturing Firms: χ^2 Measure				
χ^2 Connected	0.016	0.085**	0.082**	0.103**
	(0.019)	(0.038)	(0.035)	(0.045)
Observations	29,972	34,285	32,535	31,579
Year & Firm FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effect of political connections on firm production outcomes, as specified by equation (1). Panel A reports results for manufacturing firms and panel B reports results for manufacturing firms using the alternative χ^2 definition of political connections. All dependent variables are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C5: Difference-in-Differences Estimates: Effect of Political Connections: Excluding Controls

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Connected	0.137**	0.105**	0.404*	0.001	0.089***	0.102***	0.071
	(0.062)	(0.047)	(0.223)	(0.017)	(0.033)	(0.029)	(0.051)
Observations	64,490	82,063	82,705	29,972	87,383	73,153	50,593
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No

Notes: This table reports the estimated effect of political connections on firm outcomes, as specified by equation (1). All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C6: Difference-in-Differences Estimates: Effect of Political Connections Relative to All Non-Connected Firms

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. All Firms							
Connected	0.117***	0.120***	0.320**	0.021*	0.101***	0.090***	0.102**
	(0.042)	(0.042)	(0.130)	(0.013)	(0.033)	(0.027)	(0.042)
Panel B. All Firms: χ^2 Measure							
χ^2 Connected	0.109***	0.092***	0.357***	0.015	0.097***	0.076***	0.120***
	(0.029)	(0.023)	(0.107)	(0.017)	(0.023)	(0.019)	(0.025)
Observations	64,490	82,063	82,705	29,972	87,383	73,153	50,593
Panel C. Manufacturing Firms							
Connected	0.071*	0.104*	0.265**	0.021*	0.057	0.047	0.057
	(0.039)	(0.061)	(0.119)	(0.013)	(0.049)	(0.039)	(0.040)
Panel D. Manufacturing Firms: χ^2 Measure							
χ^2 Connected	0.101**	0.110**	0.513**	0.015	0.090**	0.068	0.104**
	(0.042)	(0.046)	(0.198)	(0.017)	(0.038)	(0.041)	(0.044)
Observations	32,228	33,482	33,558	29,972	34,285	32,535	31,579
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effect of political connections on firm outcomes relative to all non-connected firms. Panel A reports results for the full sample of firms, panel B reports results for the alternative χ^2 definition of political connections, panel C reports results from restricting the analysis to firms in the manufacturing sector only, and panel D reports results for the χ^2 definition of political connections restricting the analysis to only firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C7: Difference-in-Differences Estimates: Effect of Political Connections: Adjusting for Surname Frequency

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Weight by Connected Surname Frequency							
Connected	0.189***	0.237**	0.325	−0.013	0.192*	0.097***	0.137***
	(0.071)	(0.102)	(0.286)	(0.018)	(0.099)	(0.029)	(0.044)
Panel B. Controlling for Surname Reddy							
Connected	0.124**	0.097**	0.265	0.014	0.090***	0.081***	0.065
	(0.053)	(0.046)	(0.188)	(0.015)	(0.032)	(0.029)	(0.049)
Observations	64,490	82,063	82,705	29,972	87,383	73,153	50,593
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effect of political connections on firm outcomes using two approaches to adjust for surname frequency. Panel A shows results when weighting regressions by the inverse of a firm's connected surname's frequency. Unconnected firms are assigned the maximum of connected surname frequency to downweight their role in the estimations. Panel B shows results when controlling for the surname Reddy, which is the most common surname among connected directors. All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C8: Difference-in-Differences Estimates: Effect of Political Connections: Exact Surname Matches

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. All Firms							
Connected	0.144**	0.094**	0.146	0.033**	0.082***	0.078***	0.084
	(0.061)	(0.047)	(0.217)	(0.015)	(0.029)	(0.017)	(0.054)
Observations	64,490	82,063	82,705	29,972	87,383	73,153	50,593
Panel B. Manufacturing Firms							
Connected	0.125**	0.121**	−0.187	0.033**	0.069	0.080	0.034
	(0.061)	(0.048)	(0.346)	(0.015)	(0.048)	(0.049)	(0.045)
Observations	32,228	33,482	33,558	29,972	34,285	32,535	31,579
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effect of political connections on firm outcomes using only exact surname matches (Levenshtein distance = 0) to define connections. All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C9: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions on Firm Size

	Sales		Expenses		Profit	
	(1)	(2)	(3)	(4)	(5)	(6)
Ctd. × Post × τ_{k87}^{Input}	−0.065**	−0.062**	−0.064***	−0.057**	−0.29**	−0.34***
	(0.030)	(0.031)	(0.023)	(0.024)	(0.12)	(0.11)
Ctd. × Post × τ_{k87}^{Output}	−0.0049	−0.0020	−0.0033	−0.0000024	−0.0085	0.0066
	(0.0038)	(0.0035)	(0.0030)	(0.0031)	(0.016)	(0.016)
Ctd. × τ_{k87}^{Input}	0.069**	0.062**	0.053*	0.042	0.27**	0.32***
	(0.031)	(0.031)	(0.027)	(0.028)	(0.13)	(0.12)
Ctd. × τ_{k87}^{Output}	0.0034	0.00059	0.00099	−0.0016	0.019	0.0081
	(0.0035)	(0.0031)	(0.0028)	(0.0028)	(0.017)	(0.014)
Ctd. × Post	−0.41**	−0.46**	−0.57**	−0.64**	−1.55*	−2.15**
	(0.19)	(0.19)	(0.23)	(0.26)	(0.83)	(0.89)
Ctd.	0.51***	0.55***	0.63***	0.69**	1.63*	2.10**
	(0.19)	(0.19)	(0.24)	(0.26)	(0.83)	(0.90)
Observations	31,883	31,883	33,118	33,118	33,193	33,193
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × Industry FE	No	Yes	No	Yes	No	Yes

Notes: This table reports estimates of how tariff reductions affect the returns to political connections in terms of firm size as specified by equation (5). The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and linear industry-specific trends. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C10: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions on Firm Production

	TFP		Capital		Wage Bill		Material	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ctd. × Post × τ_{k87}^{Input}	−0.0038	−0.011	−0.044**	−0.043*	−0.047**	−0.039*	−0.067**	−0.063*
	(0.0050)	(0.0072)	(0.022)	(0.023)	(0.021)	(0.020)	(0.033)	(0.033)
Ctd. × Post × τ_{k87}^{Output}	−0.0017	−0.0018	−0.0024	−0.00010	−0.00069	0.00056	−0.0060	−0.0040
	(0.0017)	(0.0019)	(0.0039)	(0.0036)	(0.0031)	(0.0032)	(0.0039)	(0.0036)
Ctd. × τ_{k87}^{Input}	0.0045	0.011	0.045*	0.042	0.043*	0.033	0.068**	0.061*
	(0.0048)	(0.0069)	(0.026)	(0.028)	(0.022)	(0.022)	(0.034)	(0.033)
Ctd. × τ_{k87}^{Output}	0.0020	0.0020	0.0013	−0.00041	−0.00082	−0.0021	0.0045	0.0027
	(0.0014)	(0.0018)	(0.0034)	(0.0030)	(0.0032)	(0.0030)	(0.0040)	(0.0034)
Ctd. × Post	0.011	−0.063	−0.33**	−0.42***	−0.32*	−0.28	−0.55**	−0.54***
	(0.075)	(0.086)	(0.16)	(0.16)	(0.18)	(0.18)	(0.21)	(0.20)
Ctd.	0.0069	0.074	0.38**	0.47***	0.38**	0.34*	0.56**	0.54**
	(0.070)	(0.078)	(0.17)	(0.17)	(0.18)	(0.18)	(0.23)	(0.21)
Observations	29,671	29,671	33,914	33,914	32,191	32,191	31,262	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × Industry FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: This table reports estimates of how tariff reductions affect the returns to political connections in terms of firm production as specified by equation (5). The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and linear industry-specific trends. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C11: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Controlling for Delicensing and FDI Reforms

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ctd. × Post × τ_{k87}^{Input}	−0.058*	−0.062**	−0.263**	−0.009	−0.042*	−0.038*	−0.055*
	(0.030)	(0.024)	(0.110)	(0.007)	(0.023)	(0.020)	(0.030)
Ctd. × Post × τ_{k87}^{Output}	−0.003	−0.002	0.003	−0.002	−0.002	−0.000	−0.003
	(0.004)	(0.003)	(0.016)	(0.002)	(0.004)	(0.003)	(0.003)
Ctd. × Delicense	0.051	−0.236	−1.805*	0.040	−0.183	−0.004	0.008
	(0.103)	(0.253)	(1.070)	(0.099)	(0.160)	(0.098)	(0.127)
Ctd. × FDI	0.161	0.170	1.666*	−0.044	0.144	0.147	0.091
	(0.160)	(0.146)	(0.952)	(0.059)	(0.114)	(0.105)	(0.158)
Observations	25,793	26,756	26,822	23,945	27,466	26,028	25,265
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how the tariff reductions affect the returns to political connections as specified in equation (5) when controlling for whether the industry was delicensed or experienced an FDI reform at time t and their interactions with *Ctd*. The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C12: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Excluding Controls

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ctd. × Post × τ_{k87}^{Input}	−0.073**	−0.066**	−0.334***	−0.012	−0.043*	−0.051**	−0.071**
	(0.031)	(0.028)	(0.122)	(0.009)	(0.022)	(0.025)	(0.032)
Ctd. × Post × τ_{k87}^{Output}	−0.005	−0.003	0.001	−0.002	−0.003	0.001	−0.005
	(0.004)	(0.004)	(0.014)	(0.002)	(0.005)	(0.004)	(0.004)
Observations	31,883	33,118	33,193	29,671	33,914	32,191	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No

Notes: This table reports estimates of how tariff reductions affect the returns to political connections as specified in equation (5). The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C13: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Nonlinear AC Trends

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ctd. × Post × τ_{k87}^{Input}	−0.101***	−0.092***	−0.219**	−0.016**	−0.050***	−0.066***	−0.101***
	(0.024)	(0.019)	(0.106)	(0.007)	(0.014)	(0.020)	(0.027)
Ctd. × Post × τ_{k87}^{Output}	0.002	0.003	0.003	0.001	0.003	0.004	−0.004
	(0.008)	(0.006)	(0.036)	(0.004)	(0.010)	(0.006)	(0.008)
Observations	31,883	33,118	33,193	29,671	33,914	32,191	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × AC FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how tariff reductions affect returns to political connections when additionally controlling for assembly constituency-by-year fixed effects. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C14: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: χ^2-Definition

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
χ^2 Ctd. $\times$ Post $\times$ τ_{k87}^{Input}	-0.040^{*}	-0.044^{**}	-0.159^{*}	0.003	-0.027^{**}	-0.034^{**}	-0.043^{*}
	(0.022)	(0.018)	(0.093)	(0.006)	(0.011)	(0.015)	(0.023)
χ^2 Ctd. $\times$ Post $\times$ τ_{k87}^{Output}	-0.002	-0.001	-0.010	-0.000	-0.001	0.000	-0.003
	(0.003)	(0.002)	(0.017)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	31,883	33,118	33,193	29,671	33,914	32,191	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how the tariff reductions affect the returns to political connections as specified in equation (5) when using the χ^2 measure of connectedness. The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. $\chi^2 Ctd$ is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C15: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Exact Surname Matches

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ctd. $\times$ Post $\times$ τ_{k87}^{Input}	-0.087^{***}	-0.087^{***}	-0.367^{***}	-0.003	-0.058^{***}	-0.065^{***}	-0.099^{***}
	(0.028)	(0.017)	(0.098)	(0.005)	(0.017)	(0.021)	(0.028)
Ctd. $\times$ Post $\times$ τ_{k87}^{Output}	-0.006^{*}	-0.004	-0.007	-0.002	-0.003	-0.001	-0.007^{***}
	(0.003)	(0.003)	(0.016)	(0.002)	(0.003)	(0.003)	(0.003)
Observations	31,883	33,118	33,193	29,671	33,914	32,191	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how the tariff reductions affect the returns to political connections as specified in equation (5) using only exact surname matches (Levenshtein distance = 0) to define connections. The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C16: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Winsorized Data

	Sales	Expenses	Profit	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Winsorized 5% Across Sectors						
Ctd. × Post × τ_{k87}^{Input}	−0.066**	−0.061**	−0.17	−0.043**	−0.047**	−0.071**
	(0.030)	(0.023)	(0.21)	(0.020)	(0.019)	(0.034)
Ctd. × Post × τ_{k87}^{Output}	−0.0051	−0.0033	0.0058	−0.0021	−0.00053	−0.0040
	(0.0036)	(0.0029)	(0.029)	(0.0038)	(0.0030)	(0.0040)
Observations	31,883	33,139	33,193	33,915	32,200	31,514
Panel B. Winsorized 5% Within Sectors						
Ctd. × Post × τ_{k87}^{Input}	−0.061**	−0.054**	−0.34	−0.043**	−0.044*	−0.062*
	(0.030)	(0.024)	(0.33)	(0.021)	(0.022)	(0.034)
Ctd. × Post × τ_{k87}^{Output}	−0.0052	−0.0035	0.0090	−0.0022	−0.00022	−0.0051
	(0.0033)	(0.0028)	(0.044)	(0.0039)	(0.0031)	(0.0039)
Observations	31,883	33,139	33,193	33,915	32,200	31,507
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how tariff reductions affect returns to political connections after winsorizing the top and bottom 5% of the sample for each outcome. In panel A, the sample is winsorized across 2-digit NIC codes and in panel B, the sample is winsorized within 2-digit NIC codes. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C17: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Excluding Firms in the Densest ACs

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ctd. × Post × τ_{k87}^{Input}	−0.070**	−0.065***	−0.311**	−0.005	−0.045**	−0.048**	−0.071**
	(0.030)	(0.024)	(0.127)	(0.004)	(0.023)	(0.021)	(0.034)
Ctd. × Post × τ_{k87}^{Output}	−0.005	−0.003	−0.009	−0.002	−0.001	−0.000	−0.006
	(0.004)	(0.003)	(0.018)	(0.001)	(0.004)	(0.003)	(0.004)
Observations	28,199	29,272	29,345	26,224	29,971	28,472	27,659
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how the tariff reductions affect the returns to political connections as specified in equation (5) when excluding firms in assembly constituencies below the 10th percentile in terms of size (area). The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency and its interactions with tariffs and *Post*. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C18: Triple-Differences Estimates: Effect of Political Connections and Tariff Reductions: Specification using Yearly Tariffs

	Sales	Expenses	Profit	TFP	Capital	Wage Bill	Material
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Connected $\times$ τ_{kt}^{Input}	0.032***	0.026**	0.050	−0.00075	0.020**	0.021**	0.036***
	(0.0090)	(0.013)	(0.053)	(0.0033)	(0.0083)	(0.0093)	(0.011)
Connected $\times$ τ_{kt}^{Output}	−0.0021**	−0.0011	−0.0026	0.00056	−0.0012*	−0.0020**	−0.0021**
	(0.00100)	(0.0017)	(0.0062)	(0.00049)	(0.00073)	(0.00080)	(0.0010)
τ_{kt}^{Input}	−0.016**	−0.018**	−0.12***	0.000081	−0.0091	−0.012	−0.013
	(0.0080)	(0.0088)	(0.047)	(0.0038)	(0.0060)	(0.0074)	(0.0091)
τ_{kt}^{Output}	0.00061	−0.00093	0.0034	0.00034	0.000031	−0.00028	−0.00038
	(0.0012)	(0.0012)	(0.0066)	(0.00100)	(0.00094)	(0.00097)	(0.0014)
Connected	0.092	0.091	0.11	0.014	0.051	0.074	0.0096
	(0.068)	(0.061)	(0.16)	(0.015)	(0.044)	(0.046)	(0.055)
Observations	31,883	33,118	33,193	29,671	33,914	32,191	31,262
Year & Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of how the tariff reductions affect the returns to political connections when using variation in tariffs across industries and years. The specification is $y_{ikct} = \rho_1 Connected_{ict} + \rho_2 \tau_{kt}^{Input} + \rho_3 \tau_{kt}^{Output} + \rho_4 (Connected_{ict} \times \tau_{kt}^{Output}) + \rho_5 (Connected_{ict} \times \tau_{kt}^{Input}) + \alpha_i + \gamma_t + \boldsymbol{\gamma} \mathbf{X}_{ikct} + \upsilon_{ikct}$. All dependent variables, except for profit, are in logs. *Connected* is a dummy variable that takes the value of 1 if a firm is defined as connected to the politician who won the election and 0 otherwise. τ_{kt}^{Output} and τ_{kt}^{Input} denote yearly output and input tariff rates at the 4-digit NIC industry level and are demeaned. All regressions control for whether a firm is connected to either the winner or the runner-up politician within the constituency and its interactions with tariffs. Other controls are age fixed effects and a linear industry-specific (4-digit NIC code) trend. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C19: Connections, Tariff Reductions, and Sample Attrition

	Pr(Leave Sample)	
	(1)	(2)
Panel A. Connections		
Ctd.	−0.0032	−0.0036
	(0.0062)	(0.0061)
Mean of Outcome	0.03	0.03
Observations	116,890	116,890
Year & Firm FE	Yes	Yes
Controls	No	Yes
Panel B. Connections and Tariffs		
Ctd. × Post × τ_{k87}^{Input}	0.004	0.005
	(0.003)	(0.003)
Ctd. × Post × τ_{k87}^{Output}	0.0001	−0.0000
	(0.0003)	(0.0003)
Ctd.	0.07643	0.10316
	(0.1147)	(0.1305)
Mean of Outcome	0.02	0.02
Observations	44,491	44,491
Year & Firm FE	Yes	Yes
Controls	No	Yes

Notes: This table reports the results from testing whether politically connected firms are more likely to drop out of the sample. The dependent variable is a dummy equal to 1 if a firm did not report any data in year t. *Ctd* is a dummy variable equal to 1 if a firm is connected to the winner and 0 otherwise. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions control for whether a firm is connected to either the winner or runner-up politician within the constituency. Other controls are age fixed effects and industry (4-digit NIC code) by year fixed effects. Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C20: Diff-in-Disc Estimates: Effect of Political Connections and Tariff Reductions

	Sales		Expenses		Profit		Capital		Wage Bill		Material	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ctd. $\times$ Post $\times$ τ_{k87}^{Input}	-0.489^{***}	-0.707^{***}	-0.433^{***}	-0.637^{***}	-1.683^{***}	-1.475^{**}	-0.344^{***}	-0.384^{***}	-0.493^{***}	-0.469^{***}	-0.245^{**}	-0.627^{***}
	(0.102)	(0.095)	(0.092)	(0.130)	(0.359)	(0.680)	(0.051)	(0.067)	(0.144)	(0.129)	(0.111)	(0.107)
Ctd. $\times$ Post $\times$ τ_{k87}^{Output}	-0.151^{**}	-0.242^{***}	-0.136	-0.227^{**}	-0.454^{**}	-0.967^{***}	-0.024	-0.184^{***}	-0.069	-0.193	-0.115	-0.190^{**}
	(0.075)	(0.064)	(0.087)	(0.088)	(0.188)	(0.191)	(0.049)	(0.064)	(0.073)	(0.128)	(0.115)	(0.074)
Ctd. $\times$ τ_{k87}^{Input}	0.436^{***}	0.600^{***}	0.425^{***}	0.577^{***}	1.937^{***}	1.581^{**}	0.345^{***}	0.353^{***}	0.471^{***}	0.404^{***}	0.214^{**}	0.553^{***}
	(0.075)	(0.093)	(0.087)	(0.121)	(0.330)	(0.650)	(0.026)	(0.068)	(0.150)	(0.126)	(0.104)	(0.108)
Ctd. $\times$ τ_{k87}^{Output}	0.142^{*}	0.219^{***}	0.133	0.221^{**}	0.423^{**}	0.997^{***}	0.054	0.203^{***}	0.070	0.203	0.116	0.185^{**}
	(0.073)	(0.061)	(0.085)	(0.091)	(0.178)	(0.213)	(0.045)	(0.064)	(0.071)	(0.127)	(0.114)	(0.077)
Ctd. $\times$ Post	-1.426	-2.159	-1.442	-0.904	-8.785	-10.875^{*}	-0.939^{*}	-1.035	-0.419	-1.628	-0.750	-2.421
	(0.888)	(1.526)	(1.008)	(1.691)	(5.432)	(6.387)	(0.499)	(1.375)	(1.694)	(1.857)	(1.085)	(1.739)
τ_{k87}^{Input} $\times$ Post	0.204^{**}	0.594^{***}	0.160^{*}	0.487^{***}	0.882^{***}	1.902^{***}	0.122^{**}	0.392^{***}	0.147	0.344^{***}	0.041	0.539^{***}
	(0.101)	(0.092)	(0.094)	(0.124)	(0.180)	(0.355)	(0.056)	(0.065)	(0.114)	(0.089)	(0.099)	(0.103)
τ_{k87}^{Output} $\times$ Post	0.163^{**}	0.179^{***}	0.146^{*}	0.155^{*}	0.413^{***}	0.848^{***}	0.000	0.095	0.055	0.168	0.120	0.154^{**}
	(0.075)	(0.062)	(0.087)	(0.090)	(0.118)	(0.144)	(0.049)	(0.065)	(0.069)	(0.126)	(0.114)	(0.072)
Ctd.	2.053^{**}	3.003^{**}	1.561	1.470	8.832^{*}	11.487^{**}	1.413^{***}	1.587	0.420	1.665	1.174	3.240^{**}
	(0.826)	(1.268)	(0.967)	(1.467)	(5.221)	(5.215)	(0.242)	(1.190)	(1.760)	(1.682)	(1.092)	(1.551)
τ_{k87}^{Input}	-0.215^{***}	0.000	-0.189^{**}	0.000	-1.222^{***}	0.000	-0.199^{***}	0.000	-0.195^{*}	0.000	-0.032	0.000
	(0.057)	(0.000)	(0.073)	(0.000)	(0.155)	(0.000)	(0.025)	(0.000)	(0.110)	(0.000)	(0.080)	(0.000)
τ_{k87}^{Output}	-0.139^{*}	0.000	-0.125	0.000	-0.333^{***}	0.000	-0.002	0.000	-0.049	0.000	-0.104	0.000
	(0.072)	(0.000)	(0.084)	(0.000)	(0.094)	(0.000)	(0.045)	(0.000)	(0.065)	(0.000)	(0.113)	(0.000)
Post	0.704	0.000	0.844	0.000	6.255^{**}	0.000	0.291	0.000	0.559	0.000	0.436	0.000
	(0.791)	(0.000)	(0.939)	(0.000)	(2.651)	(0.000)	(0.459)	(0.000)	(1.294)	(0.000)	(0.920)	(0.000)
Observations	766	766	866	866	1,058	1,058	763	763	850	850	803	803
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Bandwidth	0.087	0.087	0.097	0.097	0.115	0.115	0.080	0.080	0.097	0.097	0.089	0.089

Notes: This table reports the estimated effect of tariff reductions on the returns to political connections using the Diff-in-Disc design (as specified in equation (6)). The sample is restricted to firms in the manufacturing sector. All dependent variables, except for profit, are in logs. *Ctd* is a dummy equal to 1 for firms connected to winning politicians and 0 for firms connected to runner-up politicians. τ_{k87}^{Output} and τ_{k87}^{Input} denote (demeaned) 1987 output and input tariff rates at the 4-digit NIC industry level. *Post* is a dummy variable equal to 1 for years after 1993 and 0 otherwise. All regressions fully interact τ_{k87}^{Input}, τ_{k87}^{Output}, and *Post* with $Margin$, Ctd, and $Margin \times Ctd$. The Diff-in-Disc estimates are the coefficients on the triple interactions $Ctd \times Post \times \tau_{k87}$. Controls are 4-digit NIC fixed effects and year fixed effects. The regressions are weighted by a triangular kernel to give more weight on observations close to the cutoff. The bandwidth is calculated using the method proposed by Calonico, Cattaneo and Titiunik (2014). Standard errors are two-way clustered by constituencies and industries. Statistical significance is indicated by *** at 1%, ** at 5%, and * at 10%.

Table C21: Summary Statistics for Close Elections Samples

	Pre-1993 CE Sample		Post-1993 CE Sample		Comparison	
	Mean (1)	Obs. (2)	Mean (3)	Obs. (4)	Diff. (5)	*p*-value (6)
Panel A. 1987–1993						
Elections and Politicians						
Incumbent	0.169	272	0.171	362	0.002	0.943
Female Candidate	0.037	272	0.044	362	0.007	0.641
Aligned Candidate	0.452	272	0.456	362	0.004	0.928
INC Candidate	0.360	272	0.362	362	0.002	0.967
BJP Candidate	0.165	272	0.149	362	−0.016	0.577
SC/ST Reserved	0.110	272	0.110	362	0.000	0.994
Contestants	12.691	272	12.210	362	−0.481	0.400
Turnout	0.589	272	0.592	362	0.003	0.781
Firms						
Age	29.399	664	28.984	933	−0.415	0.708
No. of Directors	9.110	664	9.174	933	0.064	0.735
Private, Domestic	0.974	664	0.981	933	0.006	0.396
Private, Foreign	0.018	664	0.013	933	−0.005	0.399
State Owned	0.008	664	0.006	933	−0.001	0.794
Publicly Listed	0.660	664	0.680	933	0.020	0.405
Panel B. 1994–2007						
Elections and Politicians						
Incumbent	0.235	586	0.248	840	0.012	0.600
Female Candidate	0.077	586	0.073	840	−0.004	0.768
Aligned Candidate	0.469	586	0.460	840	−0.010	0.716
INC Candidate	0.309	586	0.323	840	0.014	0.583
BJP Candidate	0.147	586	0.148	840	0.001	0.964
SC/ST Reserved	0.133	586	0.138	840	0.005	0.787
Contestants	10.590	586	10.329	840	−0.262	0.449
Turnout	0.645	586	0.645	840	−0.000	0.974
Firms						
Age	26.153	12,145	25.331	17,005	−0.822	0.001
No. of Directors	6.124	12,145	6.115	17,005	−0.009	0.828
Private, Domestic	0.943	12,145	0.942	17,005	−0.000	0.917
Private, Foreign	0.035	12,145	0.031	17,005	−0.004	0.060
State Owned	0.022	12,145	0.027	17,005	0.004	0.022
Publicly Listed	0.427	12,145	0.416	17,005	−0.011	0.072

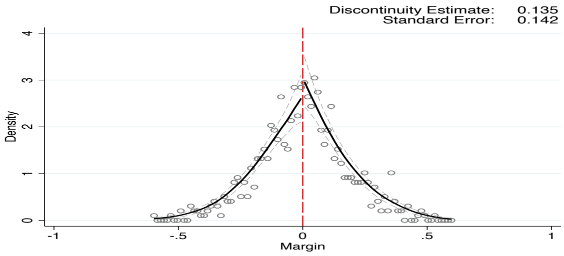
Notes: This table reports descriptive statistics for the samples of constituencies with close elections where at least one firm is defined as connected to a politician by our surname-matching methodology. To classify close elections, we use the average of the optimal bandwidths (Calonico, Cattaneo and Titiunik, 2014) for the main Diff-in-Disc outcomes: 0.098. *Pre-1993 CE Sample* refers to the sample of constituencies that, at some point during 1987–1993, had close elections and at least one connected firm. *Post-1993 CE Sample* refers to the sample of constituencies that at some point during 1994–2007 had close elections and at least one connected firm. Panel A presents statistics for the years 1987–1993 for both samples, while panel B presents statistics for the years 1993–2007. Columns (5) and (6) thus compare characteristics across the two close elections samples within the same time period, highlighting possible compositional differences in areas with close elections over time.

D RD Identification Assumptions and Results

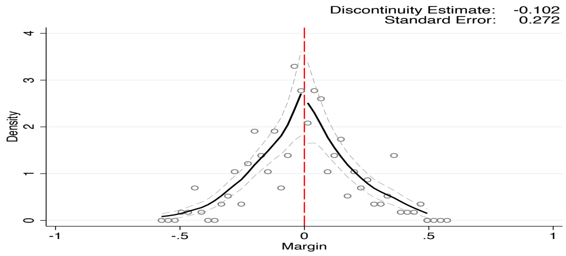
For θ_1 in equation (4) to be well identified, all confounding factors must vary smoothly at the treatment threshold. This assumption would be violated if firms with structural advantages were disproportionately connected to politicians who win close elections. To alleviate these concerns, we check the density of the running variable around the threshold and test for continuity of predetermined covariates.

Figure D1 shows the density test proposed by McCrary (2008). We find no evidence of bunching on either side of the threshold. This implies that politicians with connections to firms are not more likely to win close races. In Figure D2, we validate the continuity assumption by estimating equation (4) using predetermined variables as outcomes. For comparability reasons, we standardize each outcome variable to mean 0 and standard deviation 1. The results support that predetermined political variables, such as incumbency status, candidate gender, alignment status, political parties, constituency reservation status, number of candidates, and turnout, are balanced overall across the threshold. Moreover, we find no discontinuity in tariffs at the threshold. Finally, we investigate the continuity of pretreatment firm outcomes. We find no evidence of imbalances in these outcomes.[5] Together, these findings validate that any discontinuity in firm outcomes around the threshold can be attributed to firms gaining political connections.

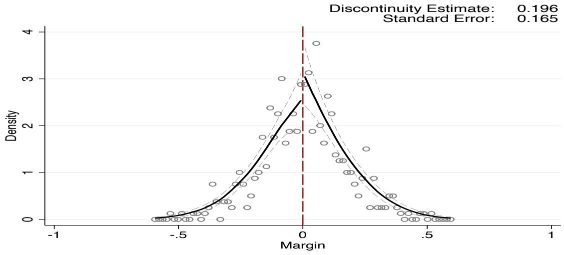
[5]A natural candidate year for determining pretreatment status is the year before an election. However, since a firm, in principle, could be treated in the preceding election, we instead use the first observed value of these variables as the pretreatment outcome.



(a) Full Period



(b) 1987–1996



(c) 1997–2007

Figure D1: Density Test for Manipulation of the Running Variable

Notes: This figure shows the density of politicians with different win-margins and plots a nonparametric regression to each half of the distribution, following McCrary (2008), to test for a discontinuity at the zero threshold. Panel (a) reports the discontinuity estimate for the full sample period. Panels (b)–(c) report estimated discontinuities for 1987–1996 and 1997–2007, respectively.

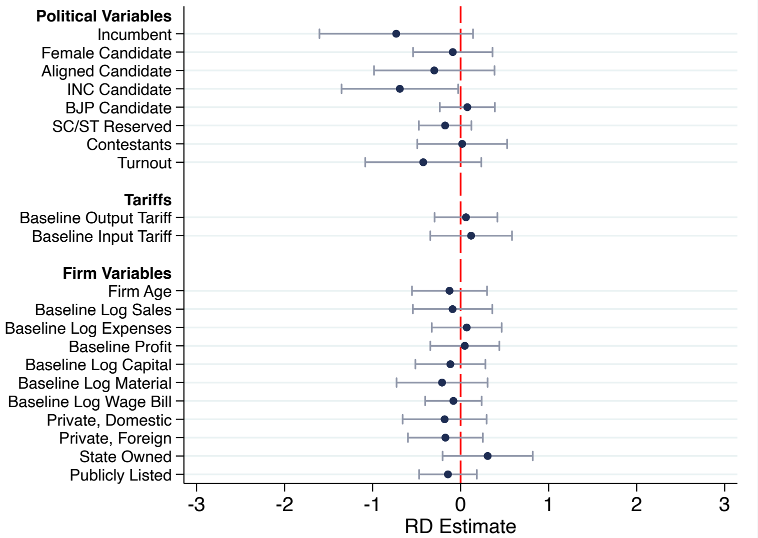
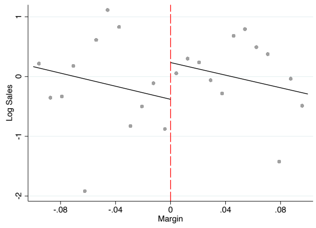
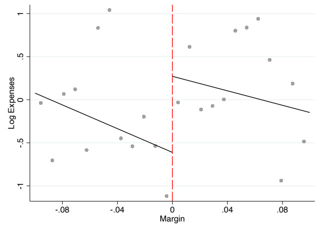
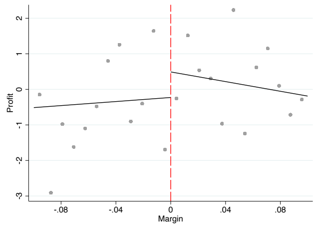


Figure D2: Continuity Test for Baseline Variables

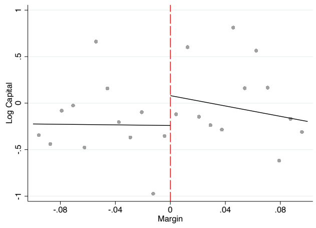
Notes: This figure shows the continuity test for pre-determined outcome variables by regressing each row variable on Connected, Margin, and their interaction, along with the set of fixed effects described in the paper. The regressions are weighted by a triangular kernel to give more weight to observations close to the threshold. All outcome variables are standardized to mean 0 and standard deviation 1 for comparability. Each baseline firm variable is predetermined in the sense that it reflects the value of the outcome variable measured at the first observation year for a firm. Regressions are estimated using the average of the optimal bandwidths (Calonico, Cattaneo and Titiunik, 2014) for the main RD outcomes: 0.095. Standard errors are two-way clustered by industry and constituency.

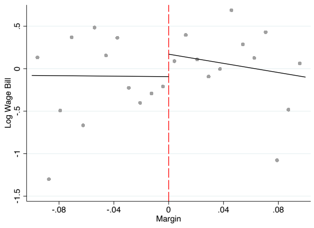


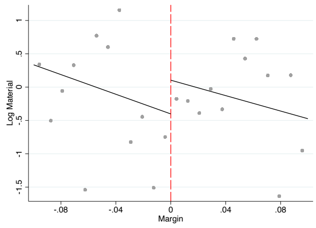




(a) Log Sales (b) Log Expenses (c) Profit



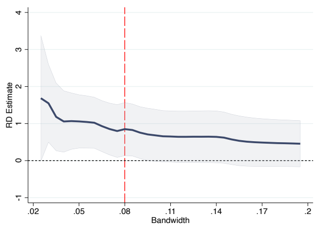


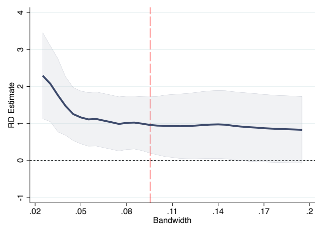


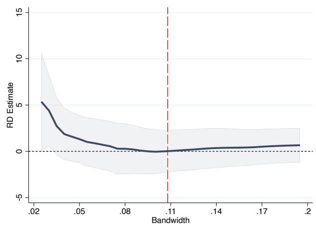
(d) Log Capital (e) Log Wage Bill (f) Log Material

Figure D3: RD Graphs: Effect of Political Connections

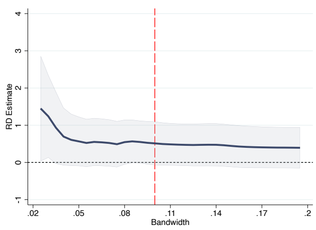
Notes: This figure displays the effect of political connections on firm outcomes by estimating a discontinuity. We partial out age, state, and year fixed effects from the outcomes before plotting these figures, in line with our robustness specification. Markers represent averages within bins. The estimation bandwidth is 10% on each side of the threshold. 95% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.

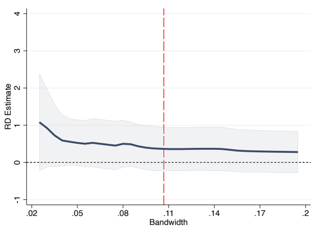


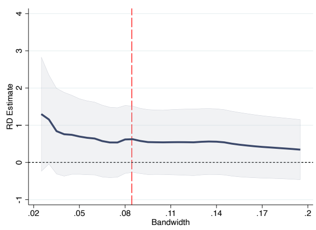




(a) Log Sales (b) Log Expenses (c) Profit







(d) Log Capital (e) Log Wage Bill (f) Log Material

Figure D4: Bandwidth Sensitivity of RD Estimates

Notes: This figure shows estimated RD coefficients for a range of different bandwidths. The specification and indexing are as described in the paper. The vertical dotted line indicates the optimal bandwidth as calculated in Calonico, Cattaneo and Titiunik (2014). Controls are age, state, and year fixed effects. 95% confidence intervals are used for inference. Standard errors are two-way clustered by industry and constituency.

References

Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad. 2021. "Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform." *World Bank Economic Review*.

Bhagavatula, Suresh, Bhalla Manaswini, Manisha Goel, and Balagopal Vissa. 2023. "Social Diversity in Corporate Boards and Firm Outcomes." *Journal of Corporate Finance*, 83.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2022. "Difference-in-Differences Estimators of Intertemporal Treatment Effects." Working Paper.

Jensenius, Francesca, and Gilles Verniers. 2017. "Studying Indian Politics with Large-Scale Data: Indian Election Data 1961–Today." *Studies in Indian Politics*, 5(2): 269–275.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–341.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142: 698–714.

ML InfoMap. 2008. "State Assembly Constituency Boundaries and Election Data, 1985-2008." *ML InfoMap Pvt. Ltd.* https://www.mlinfomap.com/map-data.php (accessed May 17, 2019).