Online Appendix

Competition and Innovation

Ana Paula Cusolito   Alvaro Garcia-Marin   William F. Maloney

A Data Appendix

A.1 The Chilean Annual Industrial Survey (ENIA)

Our main dataset is the Encuesta Nacional Industrial Anual (Annual National Industrial Survey – ENIA) for the years 1996-2007. Data for ENIA are collected annually by the Chilean National Institute of Statistics (INE), with direct participation of Chilean manufacturing plants. ENIA covers the universe of manufacturing plants with 10 or more workers, and contains detailed information on plant characteristics, such as sales, spending on inputs, employment, wages, investment, and export status. It also collects rich information at the product-level, for every product produced by each plant, reporting sales, total variable production cost, and the number of units produced and sold, which permits constructing plant-level price indexes and backing out quantities and TFPQ. Products in ENIA are defined according to the *Clasificador Unico de Productos* (CUP). This ENIA-specific product category is comparable to the 7-digit ISIC code.[1] Of the roughly 4,800 manufacturing plants tabulated per year about 20 percent are exporters, and two-thirds are small plants (less than 50 workers). Medium-sized plants (50-150 workers) and large plants (more than 150 workers) represent 20 and 12 percent, respectively.

We exclude plant-product-year observations that have zero values for total employment, demand for raw materials, sales, or product quantities. Our main analysis exploit within plant variation before and after the entry of China into the WTO in 2001. Consequently, we only consider plants observed in at least one year before and after 2001. Finally, to avoid outliers and/or misreported prices affect our results, we exclude observations where the input or the output price deviate more than five times from the average. Our final sample consists of 29,283 plant-product-year observations.

A.2 Measures of Innovation and Technological Investment in the Chilean Data

The Chilean data offer a unique opportunity to explore the effect of product market competition on innovation investments and then on to efficiency. We merge our baseline data from ENIA with the Chilean Technological Innovation Survey (EIT), which collects detailed data on a wide variety of innovatio-related investments, including (i) overall R&D expenditures, (ii) in-house R&D expenditures, (iii) investments in machinery and equipment, (iv) patents, (v) licenses, and (vi) overall

[1] For example, the wine industry (ISIC 3132) is disaggregated by CUP into 8 different categories, such as "Sparkling wine of fresh grapes", "Cider", "Chicha", and "Mosto".

investments from ENIA. The EIT is a nationally representative survey of Chilean plants, conducted by the Chilean National Statistical Agency every two to three years, and samples about 20% of the establishments in ENIA.[2] Managers retrospectively report data on technological investment for the between survey years, in practice allowing us to generate a time series from 2000-2007, except for 2002.[3]

To test whether there are systematic differences between the plants appearing in the two data sets, we run a simple regression, of each variable of interest against a dummy variable that takes the value one 1 if the plant is included in EIT and zero otherwise, controlling for sector-year fixed effects and clustering standard errors are the industry-year level. The coefficient accompanying the EIT dummy variable is interpreted as the percentage-point difference between plants in EIT and the rest of plants included in ENIA but excluded from EIT. Table A.3 shows that plants in EIT hire 105% more workers than those in ENIA (column 1). But even conditioning for size, plants in EIT show important differences. They have 36% higher sales (column 2), are 6.5% more likely to be exporters (column 3), and have 5.8 higher revenue productivity (column 4). Physical productivity and markups do not appear systematically different across samples (columns 5 and 6). Finally, plants in EIT have about a 5.6% higher profit rate (column 7).

A.3 Productivity Estimation

Productivity – the efficiency with which establishments convert inputs into outputs –, is typically measured as the log-difference between output and the contribution of inputs used in production. Detailed data on physical inputs and outputs is generally unavailable; as a result, researchers traditionally rely on revenues to proxy for establishments' physical output. This measure of productivity is known in the literature as revenue total factor productivity (TFPR), to differentiate it from the real subject of interest, where inputs and outputs are measured in terms of physical units. This last measure of productivity is known as physical total factor productivity (TFPQ).

A recent literature highlights the importance of computing total factor productivity using inputs and outputs in terms of physical units, TFPQ (see De Loecker and Goldberg, 2014, for a review). Deflating sales by industry-level price indexes leads to underestimating the production function coefficients, because more efficient plants tend to charge lower prices. Even if the coefficients were known, the resulting TFPR measure would be an imperfect proxy of TFPQ, because the price component downward biases the response of plant revenues. Similarly, deflating materials' costs by industry-level input indexes also biases the production function coefficients. Plants facing high input prices have higher input expenditure that is not necessarily reflected in higher output (De

[2]However, all entities representing more than 2% of sectoral value-added enter compulsory in the innovation survey.

[3]We do not use earlier waves of the survey because they do not allow us to establish a one-to-one correspondence with the questions from the most recent versions of EIT.

Loecker and Goldberg, 2014). Thus, input price variation may also lead to underestimate the variation in physical total factor productivity.

The presence of multi-product plants introduces an additional challenge in the estimation of TFPQ. The use of inputs is typically not disaggregated for individual products, making the identification of the production function coefficients for each production line within plants challenging. The early literature addressed the problem eliminating multi-product plants from the sample, and focusing only on the subset of single-product plants (see Foster, Haltiwanger, and Syverson, 2008). This approach has pitfalls of its own, as multi-product plants account for a non-trivial fraction of output in the manufacturing sector. A more recent approach relies on the use of single product units, where no assumptions about allocation are needed, to estimate the production function coefficients for each product category (De Loecker et al., 2016). These are then used to infer the allocation of inputs across outputs and generate a measure of plant-level TFPQ using the algorithm developed by De Loecker et al. (2016). A drawback of this approach is that requires assuming that the technology used to produce a given product in single and multi-product units is the same, and in practice, the algorithm may yield corner solutions for inputs use across outputs.

In this paper, we follow a third approach, which consists of deriving plant-level output and input price indexes to deflate plant's revenues and materials expenditure, respectively (see Eslava et al., 2013; Smeets and Warzynski, 2013; Eslava and Haltiwanger, 2020).[4] To derive the price indexes, we perform the following steps. First, we compute for each output and input at the plant-product-year level, the log difference of its price relative to the average industrial price for the same year. Second, we construct a weighted average price deviation index, using plant-product revenues and expenditure shares, respectively, as weights. Finally, we compute the plant-level output and input indexes adding the plant-level log-deviations derived in the previous step to the average price index defined for each 4-digit ISIC sector.

Once inputs and outputs are deflated with plant-level price indexes, we estimate Cobb-Douglas production functions for each 2-digit manufacturing sector (s) using labor (l), capital (k), and materials (m) as production inputs:[5]

$$q_{it} \quad = \quad \beta_l^s l_{it} + \beta_k^s k_{it} + \beta_m^s m_{it} + \omega_{it} + \varepsilon_{it} \tag{A.1}$$

[4] A shortcoming of this more aggregate approach is that plant-level output price indexes may not account for differences in product scope Hottman, Redding, and Weinstein (2016). Though in principle the De Loecker et al.'s (2016) methodology could give us plant-product level TFPQ, our information on productivity-enhancing activities, such as R&D expenditure, is only available at the plant level. In addition, when implementing De Loecker et al.'s (2016) methodology to identify product-specific inputs use and plant-product level TFPQ, we obtained a large number of corner solutions for the use of intermediate input. This complicates the computation of plant-level markups, because markups for products with zero use of the production factors would be indeterminate.

[5] The 2-digit product categories are: Food and Beverages, Textiles, Apparel, Wood, Paper, Chemicals, Plastic, Non-Metallic Manufactures, Basic and Fabricated Metals, and Machinery and Equipment.

where all lowercase variables are expressed in logs; q_{it} stands for output of plant i in year t, l_{it} stands for labor, k_{it} stands for capital, m_{it} stands for material inputs, ω_{it} is the productivity measure, and ε_{it} captures measurement error and output shocks.

To estimate (A.1), we follow the methodology by Ackerberg et al. (2015, henceforth ACF), who extend the framework of Olley and Pakes (1996, henceforth OP) and Levinsohn and Petrin (2003, henceforth LP) to control for potential simultaneity bias that arises because input demand and unobserved productivity are positively correlated.[6] Further, we modify the canonical ACF procedure by specifying an endogenous productivity process that accounts for learning by exporting and investment effects, and we also include interaction terms between export status and investment as in De Loecker (2013) to avoid overestimating the capital coefficient and underestimating productivity. Accordingly, the productivity law of motion is as follows:

$$\omega_{it} = g(\omega_{it-1}, d^x_{it-1}, d^i_{it-1}, d^x_{it-1} \times d^i_{it-1}) + \xi_{it} \tag{A.2}$$

where d^x_{it} is an export dummy, and d^i_{it} is a dummy for periods in which a plant invests in physical capital (following De Loecker, 2013).

A.4 Estimating Plant-Level Markups

We follow De Loecker and Warzynski (2012) in deriving markups from the first order condition (FOC) of a cost minimization problem of the plant. Assuming that (i) at least one input is fully flexible, and (ii) plants minimize costs for each product j permits re-arranging the FOC for the flexible input V and obtaining the following expression for the markup of plant i at time t:

$$\underbrace{\mu_{it}}_{\text{Markup}} \equiv \frac{P_{it}}{MC_{it}} = \underbrace{\left(\frac{\partial Q_{it}(\cdot)}{\partial V_{it}} \frac{V_{it}}{Q_{it}}\right)}_{\text{Output Elasticity}} \Big/ \underbrace{\left(\frac{P^V_{it} \cdot V_{it}}{P_{it} \cdot Q_{it}}\right)}_{\text{Expenditure Share}} \tag{A.3}$$

where P (P^V) denotes the price of output Q (input V), and MC is marginal cost. This implies that the markup can be computed by dividing the output elasticity of product j with respect to the flexible input, in our case materials (M), by its expenditure share (relative to the sales of product j). Under perfect competition, the output elasticity equals the expenditure share, so that the markup is equal to one.

A.5 Inferring Product Quality

The estimation of product quality follows Khandelwal et al. (2013) and Eslava et al. (2013). Specifically, we assume that firm's i product quality in period t, λ_{it}, acts as a demand shifter in consumer

[6] We follow LP in using material inputs to control for the correlation between input levels and unobserved productivity.

preferences. Assuming CES demand, the utility function becomes:

$$U_t = \left(\int (\lambda_{it} q_{it})^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \tag{A.4}$$

implying a demand of the form $q_i = \lambda_i^{\sigma-1} p^{-\sigma} P^{\sigma-1} Y$, where $\sigma > 1$ denotes the elasticity of substitution, P is the aggregate CES price index and Y is aggregate consumption. After taking logs, the quality for the product produced by each firm can be recovered as the residual term from the following linear regression:

$$\ln q_{it} = -\sigma \ln p_{it} + \alpha_j + \alpha_y + \varepsilon_{it} \tag{A.5}$$

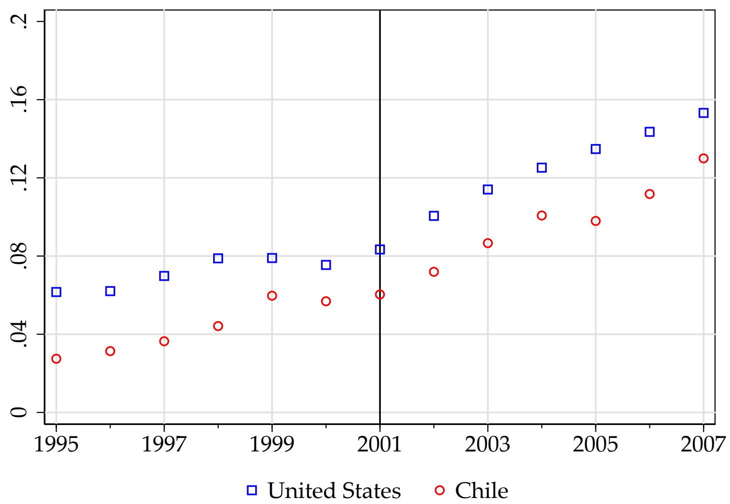
Thus, quality is constructed as a demand residual assuming a constant industry-wide price elasticity of demand. This is not necessarily inconsistent with the observed markup variation across firms within sectors, which may suggest firm-specific price elasticities of demand. In Peters (2020) firms' markups are determined by Bertrand competition with limit pricing –the price that prevents new competitors from entering– and hence the follower firm does not produce. Firms enter at the limit price but then invest to become more productive, escape competition and raise the markup. Hence, an industry wide elasticity of demand can yield distribution of markups across firms.

Estimating A.5 through OLS leads to upward bias in the estimated demand elasticity, as the quality shifter is positively correlated to input and output prices. To address this concern, we follow Eslava et al. (2013) and instrument prices using physical total factor productivity (TFPQ). Equation A.5 is estimated separately for each 3-digit manufacturing sector. Once we estimate the demand elasticity, we compute product quality as follows:

$$\widehat{\lambda_{it}} = \widehat{\alpha_j} + \widehat{\alpha_y} + \widehat{\varepsilon_{it}} \tag{A.6}$$

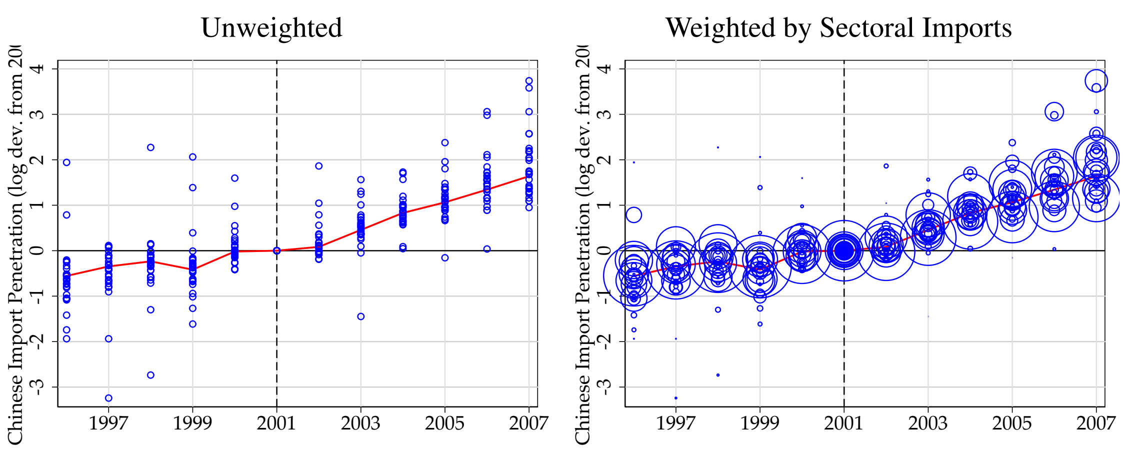
B Additional Figures

Figure A.1: Imports from China as a share of overall manufacturing imports in Chile and the U. S.



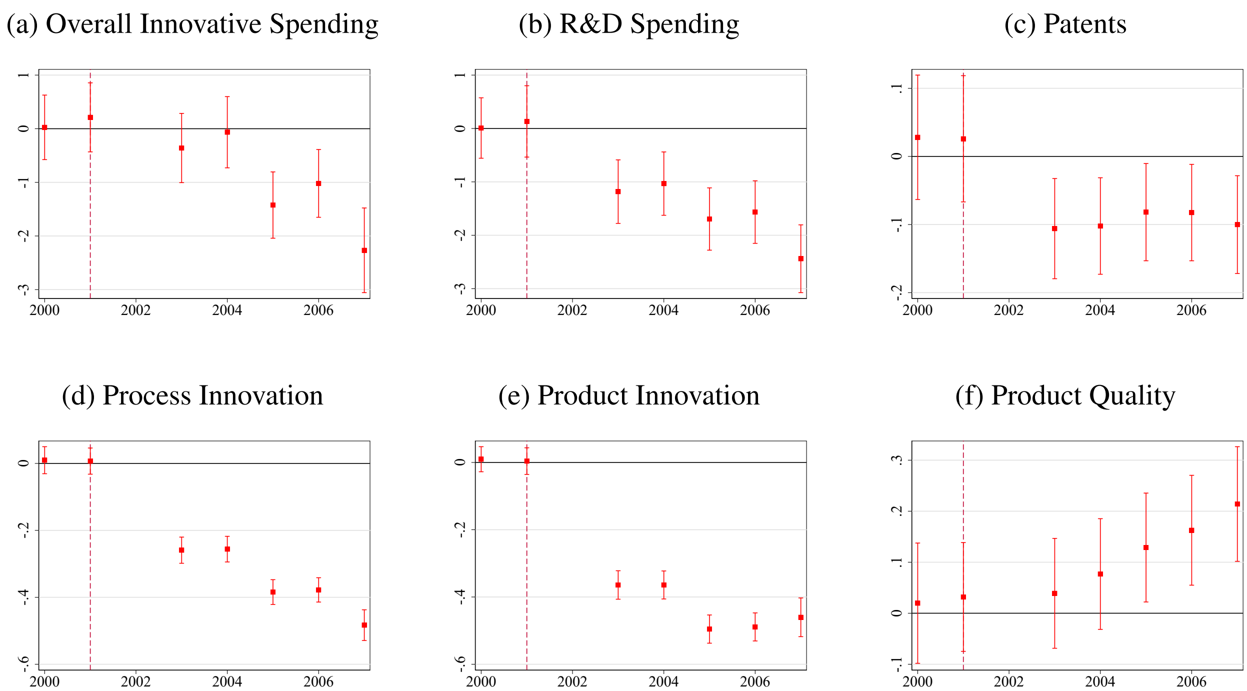
Notes: The figure shows the share of overall manufacturing imports from China in Chile and the United States for each year between 1996 and 2007. The data is from the BACI dataset (Gaulier and Zignago, 2010).

Figure A.2: Imports from China in Chile by 3-digit ISIC Sectors, 1996-2007



Notes: The figure shows the imports of Chinese products in Chile (upper panel) and the United States (bottom panel) for each 3-digit ISIC manufacturing industry (revision 2) over the period 1996-2007. The data is from the BACI dataset (Gaulier and Zignago, 2010). For each country and sector, we normalize imports equal to zero in 2001, corresponding to the year when China joined the World Trade Organization. The red line corresponds to the trajectory of aggregate manufacturing imports across all sectors.

Figure A.3: Within-Plant Trajectories for Different Innovation Outcomes



Notes: Data are from the third, fourth, fifth and sixth waves of the Chilean Technological Innovation Survey (EIT). The figure shows within-plant trajectories for different innovation outcomes before and after 2001, corresponding to the year when China joined the World Trade Organization. All results are at the plant level; they control for plant fixed effects and 2-digit sector-year fixed effects. Standard errors (clustered at the 3-digit sector-year level) in parentheses. The lines and whiskers represent 90% confidence intervals.

C Additional Tables

Table A.1: Chilean Imports of Chinese Products by 3-digit ISIC industry (1996-2007)

ISIC, Rev. 2	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
311	1.6	2.3	2.7	2.3	2.9	2.6	2.8	4.7	5.8	7.8	8.9	21.2
312	0.3	0.2	0.2	0.6	0.7	0.5	0.9	1.6	2.1	2.5	4.7	8.3
313	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
321	56.4	58.8	61.3	58.5	89.3	70.0	79.8	116.6	177.8	224.5	302.2	360.9
322	216.3	245.2	287.7	263.9	347.2	374.9	365.5	499.6	709.1	783.6	986.9	1,193.8
323	16.5	22.9	26.4	22.5	29.6	29.1	32.4	42.0	63.7	73.7	86.5	113.8
324	50.9	62.6	73.2	45.1	73.0	85.0	89.3	121.1	157.4	170.4	207.2	254.6
331	2.1	2.8	3.0	2.6	3.7	4.2	6.8	8.1	10.6	12.3	16.7	20.6
332	3.7	6.4	8.6	6.7	10.6	9.7	11.5	16.5	25.6	30.7	53.8	85.1
341	0.7	1.1	1.2	1.7	3.5	2.5	2.3	3.2	5.0	5.9	8.7	26.6
342	1.3	2.7	2.8	2.4	2.4	2.5	2.3	4.8	4.3	5.9	8.9	11.5
351	20.1	22.8	25.6	27.3	35.4	39.5	43.6	54.6	86.0	133.0	175.6	307.2
352	7.9	7.2	8.8	6.7	10.6	14.2	17.2	28.1	40.4	50.5	62.3	76.2
353	0.0	0.0	0.2	1.1	0.7	0.5	1.8	1.4	0.8	0.9	1.5	1.1
354	0.0	1.2	2.0	2.0	2.1	2.7	2.6	2.6	4.8	16.5	14.9	8.1
355	1.7	2.9	2.1	2.3	4.9	7.3	10.3	20.2	39.6	51.1	56.8	94.4
356	38.0	53.8	62.2	51.7	74.4	79.3	84.0	121.8	167.9	207.5	252.3	291.1
361	14.2	12.4	16.4	14.8	22.1	18.2	16.7	26.1	34.8	35.2	44.1	46.9
362	2.5	4.0	4.9	3.7	6.4	6.9	7.7	12.2	16.5	21.2	26.2	36.5
369	0.6	1.1	1.1	1.3	1.7	2.1	2.9	4.2	8.2	12.2	21.6	39.7
371	17.7	19.2	17.7	16.5	16.5	17.8	18.2	20.0	23.8	38.2	133.5	264.4
372	0.9	0.5	0.6	1.5	3.6	5.6	5.0	6.3	7.6	13.4	29.3	49.6
381	38.4	50.5	56.1	47.6	64.3	57.7	69.1	101.7	148.0	184.2	235.1	306.0
382	42.0	60.0	67.0	71.0	108.6	121.0	135.7	175.2	268.0	447.0	590.1	893.0
383	76.0	108.8	112.9	101.6	145.4	157.8	205.7	325.8	483.5	640.0	821.6	1,111.7
384	10.8	10.9	14.0	17.7	33.2	27.5	38.3	48.8	73.0	99.3	152.1	211.0
385	10.0	12.9	15.1	13.1	28.7	28.6	27.7	41.7	64.9	73.8	108.4	101.3
390	81.6	110.3	116.7	92.5	109.0	100.4	96.9	167.4	219.7	245.0	315.2	389.2
Overall	712.1	883.6	990.8	878.9	1,230.5	1,268.1	1,377.1	1,976.2	2,848.8	3,586.2	4,725.2	6,323.7

Notes: This table shows Chilean imports of Chinese products for each 3-digit ISIC industry (revision 2). All amounts are in million U.S. dollars of 2010.

Table A.2: Input.Output Elasticities, Returns to Scale and Markups by Sector

	Average Elasticities			Returns	Average
	Materials	Capital	Labor	to Scale	Markups
Food and Beverages	0.624	0.135	0.393	1.153	1.241
	(0.036)	(0.059)	(0.071)	(0.035)	(0.554)
Textiles	0.431	0.290	0.313	1.034	1.042
	(0.131)	(0.253)	(0.098)	(0.080)	(0.788)
Apparel and leather	0.589	0.139	0.292	1.019	1.318
	(0.081)	(0.092)	(0.144)	(0.099)	(0.772)
Wood and Furniture	0.567	0.131	0.364	1.062	1.221
	(0.073)	(0.057)	(0.073)	(0.103)	(0.71)
Paper and Printing	0.473	0.234	0.428	1.135	1.310
	(0.058)	(0.025)	(0.15)	(0.091)	(0.741)
Petroleum and Chemical industries	0.565	0.130	0.479	1.174	1.36
	(0.091)	(0.130)	(0.246)	(0.122)	(0.824)
Plastic and rubber	0.532	0.196	0.474	1.202	1.190
	(0.094)	(0.091)	(0.149)	(0.077)	(0.649)
Non-metallic products	0.640	0.119	0.402	1.162	1.603
	(0.124)	(0.104)	(0.135)	(0.132)	(0.834)
Metallic Products	0.418	0.263	0.357	1.038	1.008
	(0.038)	(0.138)	(0.141)	(0.067)	(0.671)
Machinery and Equipment	0.489	0.191	0.443	1.122	1.203
	(0.086)	(0.130)	(0.197)	(0.129)	(0.692)
Average	0.555	0.171	0.389	1.115	1.231
	(0.105)	(0.119)	(0.139)	(0.102)	(0.689)

Notes: The table shows the average elasticities, returns to scale, and markups for each 2-digit sector. Elasticities are computed as the marginal change in physical output as the input use changes. The underlying production function is Translog, and considers physical volume as output measure (deflated with plant-specific input and output price indexes, explained in section A.3). Labor is measured in terms of number of employees, materials considers expenditure deflated with a plant-specific input price index, an capital is constructed using the method of perpetual inventories. To compute markups, we follow De Loecker and Warzynski (2012), considering materials as the relevant flexible input.

Table A.3: Sample Differences: ENIA vs EIT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Plant Size		Export Prob.	Productivity		Rents	
Dependent Variable:	ln(workers)	ln(sales)	D(Exp=1)	ln(TFPR)	ln(TFPQ)	ln(markup)	Profit Rate
EIT dummy	1.047***	0.364***	0.065***	0.058***	0.025	-0.017*	0.060***
	(0.069)	(0.036)	(0.010)	(0.013)	(0.017)	(0.010)	(0.014)
Sector-year FE	✓	✓	✓	✓	✓	✓	✓
Observations	29,283	29,283	29,283	29,283	29,283	29,283	29,283

Notes: The table regresses each column variable on a categorical variable that takes the value one for observations included in the Survey of Technological Innovation (EIT). Thus, the coefficient in each column reports the log-point difference of the dependent variable between plants-years included in EIT compared to plants in the Manufacturing Survey (ENIA). All regressions control for sector-year effects at the 2-digit level and log employment – except for column 1, that only controls for sector-year fixed effects. Standard errors are clustered at the sector-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.4: First Stage Regressions, Table 2

Specification	Baseline	Leaders vs. Laggards	
	(1)	(2)	(3)
Dependent Variable:	ln($M^{CHN}_{j,t-1}$)	ln($M^{CHN}_{j,t-1}$) $\times$ Leader	ln($M^{CHN}_{j,t-1}$) $\times$ Laggard
ln($\hat{M}^{LASSO}_{j,t-1}$)	1.251***	—	—
	(0.145)		
ln($\hat{M}^{LASSO}_{j,t-1}$) $\times$ Leader	—	1.319***	—
		(0.119)	
ln($\hat{M}^{LASSO}_{j,t-1}$) $\times$ Laggard	—	—	1.238***
			(0.147)
First Stage F-statistic	74.3	35.5	
Industry-year FE	yes	yes	yes
Plant FE	yes	yes	yes
Observations	29,283	29,283	29,283

Notes: The table show first-stage regressions (equation 2 in the main text) for the IV specifications in table 2. Column 1 shows the first stage for panel A in table 2, while columns 2 and 3 shows the first stage for panel B, where (lagged) imports from China are interacted with indicator variables for leading and laggard establishments. The regressions instrument Chilean imports from China at the 3-digit level (and their interactions) with the predicted values from a LASSO regression that considers imports from China by all other countries available in the BACI trade dataset (and their interactions). Industry leaders correspond to the top 10 percent of plants with the highest average TFPQ (within industries) before 2001. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.5: Effect of Chinese Import Competition: Complete Set of Plant's Outcomes

	(1) Output	(2) Markup	(3) Revenue	(4) Output Price	(5) TFPQ	(6) TFPR	(7) Marginal Cost	(8) Quality	(9) Profits	(10) Input Price
A. Baseline										
ln(CHN Imports(-1))	-0.0441**	-0.0346***	0.0169	0.0611***	-0.0377	-0.0158	0.0973***	0.0558	-0.0554***	0.0315
	(0.0216)	(0.0108)	(0.0245)	(0.0226)	(0.0241)	(0.0134)	(0.0261)	(0.0344)	(0.0136)	(0.0240)
First-Stage F-Statistic	74.3	74.3	74.3	74.3	74.3	74.3	74.3	58.2	74.3	74.3
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	29,283	29,283	29,283	29,283	29,283	29,283	29,283	24,439	29,283	29,283
B. Interactions with Leaders / Laggards										
ln(CHN Imports(-1))										
× Leaders Indicator	0.0105	-0.102***	0.116***	0.105***	-0.0866**	-0.0445	0.220***	0.167***	-0.0917**	0.0505
	(0.0355)	(0.0224)	(0.0350)	(0.0342)	(0.0403)	(0.0322)	(0.0454)	(0.0488)	(0.0390)	(0.0326)
× Laggards Indicator	-0.0642**	-0.0249**	-0.00785	0.0564**	-0.0397	-0.0165	0.0826***	0.0369	-0.0511***	0.0286
	(0.0289)	(0.0115)	(0.0240)	(0.0232)	(0.0258)	(0.0140)	(0.0261)	(0.0314)	(0.0143)	(0.0249)
First-Stage F-Statistic	35.5	35.5	35.5	35.5	35.5	35.5	35.5	27.5	35.5	35.5
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	29,283	29,283	29,283	29,283	29,283	29,283	29,283	24,439	29283	29,283

Notes: The table presents the results from estimating equation (3) (panel A) and an extended version that interacts lagged imports from China with an indicator variable for industry leaders and laggards (panel B). Industry leaders correspond to the top 10 percent of plants with the highest average TFPQ before 2001. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. Each column shows 2SLS coefficients using (lagged) predicted LASSO imports as an instrument for (lagged) Chinese imports. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). Section 3.2 explains the procedure followed to derive the product quality measure. All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.6: Effect of Chinese Import Competition on Plants' Outcomes, Using Alternative Instrument

	(1) Output	(2) Markup	(3) Revenue	(4) Output Price	(5) TFPQ	(6) TFPR	(7) Marginal Cost	(8) Quality	(9) Profits	(10) Input Price
A. Baseline										
ln(CHN Imports(-1))	-0.033	-0.023	0.048	0.080***	-0.070**	-0.012	0.112***	0.091**	-0.038*	-0.008
	(0.029)	(0.017)	(0.033)	(0.028)	(0.031)	(0.018)	(0.034)	(0.044)	(0.021)	(0.028)
First-Stage F-Statistic	104.6	104.6	104.6	104.6	104.6	104.6	104.6	82.2	104.6	104.6
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	29,283	29,283	29,283	29,283	29,283	29,283	29,283	24,439	29,283	29,283
B. Interactions with Leaders / Laggards										
ln(CHN Imports(-1))										
$\times$ Leader Indicator	0.014	-0.071**	0.118***	0.104**	-0.091*	-0.051	0.185***	0.165***	-0.073	0.030
	(0.044)	(0.030)	(0.043)	(0.046)	(0.052)	(0.042)	(0.054)	(0.059)	(0.052)	(0.040)
$\times$ Laggard Indicator	-0.038	-0.016	0.029	0.068**	-0.060*	-0.007	0.092***	0.074*	-0.030	-0.010
	(0.039)	(0.018)	(0.033)	(0.029)	(0.034)	(0.020)	(0.034)	(0.041)	(0.022)	(0.030)
First-Stage F-Statistic	47.9	47.9	47.9	47.9	47.9	47.9	47.9	37.6	47.9	47.9
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	29,283	29,283	29,283	29,283	29,283	29,283	29,283	24,439	29283	29,283

Notes: The table replicates table 2 excluding South American countries from the computation of the LASSO instrument introduced in section 2.2. Panel A presents the results from estimating equation (3), while panel B presents results from an extended version that interacts lagged imports from China with an indicator variable for industry leaders and laggards (panel B). Industry leaders correspond to the top 10 percent of plants with the highest average TFPQ before 2001. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. Each column shows 2SLS coefficients instrumenting Chilean imports from China at the 3-digit level (and their interactions) with the predicted values from a LASSO regression that considers imports from China by all countries from outside South America available in the BACI trade dataset (and their interactions). The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). Section 3.2 explains the procedure followed to derive the product quality measure. All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.7: First Stage Regressions, Table A.6

Specification	Baseline	Leaders vs. Laggards	
	(1)	(2)	(3)
Dependent Variable:	ln($M^{CHN}_{j,t-1}$)	ln($M^{CHN}_{j,t-1}$) × Leader	ln($M^{CHN}_{j,t-1}$) × Laggard
ln($\hat{M}^{LASSO,2}_{j,t-1}$)	1.105***	—	—
	(0.108)		
ln($\hat{M}^{LASSO,2}_{j,t-1}$) × Leader	—	1.224***	—
		(0.0938)	
ln($\hat{M}^{LASSO,2}_{j,t-1}$) × Laggard	—	—	1.084***
			(0.111)
First Stage F-statistic	104.6	47.9	
Industry-year FE	yes	yes	yes
Plant FE	yes	yes	yes
Observations	29,283	29,283	29,283

Notes: The table show first-stage regressions (equation 2 in the main text) for the IV specifications in table A.6. Column 1 shows the first stage for panel A in table A.6, while columns 2 and 3 shows the first stage for panel B, where (lagged) imports from China are interacted with indicator variables for leading and laggard establishments. The regressions instrument Chilean imports from China at the 3-digit level (and their interactions) with the predicted values from a LASSO regression that considers imports from China by all countries from outside South America available in the BACI trade dataset (and their interactions). Industry leaders comprise the top 10 percent of plants with the highest average TFPQ (within industries) before 2001. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.8: Effect of Chinese Import Competition on Plants' Outcomes - EIT Sample

	(1) Output	(2) Markup	(3) Revenue	(4) Output Price	(5) TFPQ	(6) TFPR	(7) Marginal Cost	(8) Quality	(9) Profits	(10) Input Price
A. Baseline										
ln(CHN Imports(-1))	-0.031	-0.052*	0.072	0.104**	-0.099*	-0.061	0.167***	0.079	-0.099**	-0.026
	(0.040)	(0.030)	(0.049)	(0.040)	(0.056)	(0.040)	(0.057)	(0.059)	(0.050)	(0.037)
First-Stage F-Statistic	42.9	42.2	42.9	42.9	42.2	42.9	42.2	45.8	42.9	42.1
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	4,704	4,606	4,704	4,704	4,606	4,702	4,606	4,345	4,704	4,525
B. Interactions with Leaders / Laggards										
ln(CHN Imports(-1))										
× Leaders Indicator	0.022	-0.072	0.192**	0.170**	-0.166	-0.129	0.215*	0.217***	-0.044	0.026
	(0.104)	(0.060)	(0.077)	(0.066)	(0.130)	(0.103)	(0.121)	(0.077)	(0.085)	(0.082)
× Laggards Indicator	-0.035	-0.041	0.028	0.063	-0.056	-0.048	0.123*	0.023	-0.096*	-0.038
	(0.050)	(0.034)	(0.053)	(0.044)	(0.057)	(0.040)	(0.063)	(0.061)	(0.057)	(0.040)
First-Stage F-Statistic	14.5	14.4	14.5	14.5	14.4	14.5	14.4	16.00	14.5	14.30
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	4,704	4,606	4,704	4,704	4,606	4,702	4,606	4,345	4,704	4,525

Notes: The table replicates table 2 for the sample of plant-years available in the Chilean innovation survey (EIT), regressing different plant outcomes on lagged imports from China. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. Each column shows 2SLS coefficients using (lagged) predicted LASSO imports as an instrument for (lagged) Chinese imports. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). Section 3.2 explains the procedure followed to derive the product quality measure. All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.9: First Stage Regressions, Tables 3 and 4

Specification	Baseline	Leaders vs. Laggards		Change in Markups	
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	$\ln(M^{CHN}_{j,t-1})$	$\ln(M^{CHN}_{j,t-1})$ $\times$ Leader	$\ln(M^{CHN}_{j,t-1})$ $\times$ Laggard	$\ln(M^{CHN}_{j,t-1}) \times \mathbb{I}(\Delta\mu^{after}_{before} > 0)$	$\ln(M^{CHN}_{j,t-1}) \times \mathbb{I}(\Delta\mu^{after}_{before} < 0)$
$\ln(\hat{M}^{LASSO}_{j,t-1})$	1.083***	—	—	—	—
	(0.165)				
$\ln(\hat{M}^{LASSO}_{j,t-1}) \times$ Leader	—	1.304***	—	—	—
		(0.116)			
$\ln(\hat{M}^{LASSO}_{j,t-1}) \times$ Laggard	—	—	1.027***	—	—
			(0.191)		
$\ln(\hat{M}^{LASSO}_{j,t-1}) \times \mathbb{I}(\Delta\mu^{after}_{before} > 0)$	—	—	—	1.135***	—
				(0.212)	
$\ln(\hat{M}^{LASSO}_{j,t-1}) \times \mathbb{I}(\Delta\mu^{after}_{before} < 0)$	—	—	—	—	1.057***
					(0.154)
First Stage F-statistic	42.9	14.5		23.5	
Industry-year FE	yes	yes	yes	yes	yes
Plant FE	yes	yes	yes	yes	yes
Observations	4,704	4,704	4,704	4,692	4,692

Notes: The table show first-stage regressions for the IV specifications (equation 2 in the main text) in table 3 and 4. Column 1 shows the first stage for panel A in table 3, while columns 2 and 3 shows the first stage for panel B, where (lagged) imports from China are interacted with indicator variables for leading and laggard establishments. Columns 5 and 6 shows the first stage for table 4, where (lagged) imports from China are interacted with indicator variables for plants with higher/lower markups in 2001-2007 relative to the precedent period 1996-2000. The regressions instrument Chilean imports from China at the 3-digit level (and their interactions) with the predicted values from a LASSO regression that considers imports from China by all other countries available in the BACI trade dataset (and their interactions). Industry leaders correspond to the top 10 percent of plants with the highest average TFPQ (within industries) before 2001. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.10: Effect of Chinese Import Competition on Innovation Variables: Single-Plant Firms

	(1)	(2)	(3)	(4)	(5)	(6)
	Innovative Spending		Innovation Outputs			
	Overall Spending	R&D Spending	Patents Stock	Process Innovation	Product Innovation	Product Quality
A. Baseline						
ln(CHN Imports(-1))	-0.840*	-0.901*	-0.026	-0.085**	-0.064***	0.027
	(0.478)	(0.538)	(0.057)	(0.034)	(0.024)	(0.063)
Avg. elasticity (IHS variables)	*-0.840*	*-0.901*	*-0.044*	—	—	—
First-Stage F-Statistic	39.5	39.5	39.5	39.5	39.5	34.9
Industry FE	yes	yes	yes	yes	yes	yes
Observations	2,811	2,811	2,811	2,811	2,811	2,564
B. Interactions with Leaders / Laggards						
ln(CHN Imports(-1))						
× Leaders Indicator	-0.971	0.379	-0.225	-0.104	-0.003	0.094
	(1.007)	(0.608)	(0.169)	(0.078)	(0.084)	(0.057)
Avg. elasticity (IHS variables)	*-0.971*	*0.379*	*-0.235*	—	—	—
× Laggards Indicator	-0.988	-1.250*	0.010	-0.098**	-0.077***	-0.014
	(0.740)	(0.683)	(0.049)	(0.043)	(0.028)	(0.076)
Avg. elasticity (IHS variables)	*-0.988*	*-1.250*	*-0.022*	—	—	—
First-Stage F-Statistic	13.4	13.4	13.4	13.4	13.4	11.9
Industry FE	yes	yes	yes	yes	yes	yes
Observations	2,811	2,811	2,811	2,811	2,811	2,564

Notes: The table replicates table 3 for the sample of single-plant firms from the matched ENIA-EIT data. All regressions are run at the plant-year level, control for the logarithm of employment, and include industry-year (at the 2-digit level) and plant fixed-effects. Each column shows 2SLS coefficients using (lagged) predicted LASSO imports as an instrument for (lagged) Chinese imports. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). Overall innovative spending (column 1), R&D spending (column 2), and patents stock (column 3) are transformed using the inverse hyperbolic sine function (IHS, see Bellemare and Wichman, 2020) to account for zeros. Product and process innovation (columns 4-5) are categorical variables taking the value one if the establishment reports successful innovation. Section 3.2 explains the procedure followed to derive the product quality measure. All regressions cluster standard errors at the industry-year level. Key: ** significant at 1%; ** 5%; * 10%.

Table A.11: Effect on Innovation Variables: Specification in Differences

	(1)	(2)	(3)	(4)	(5)	(6)
	Innovative Spending		Innovation Outputs			
	Overall Spending	R&D Spending	Patents Stock	Process Innovation	Product Innovation	Product Quality
A. Baseline						
Δln(CHN Imports(-1))	-2.749**	-1.601*	-0.0631	-0.103*	-0.232***	0.044
	(1.013)	(0.916)	(0.149)	(0.052)	(0.051)	(0.115)
Avg. elasticity (IHS variables)	*-2.749*	*-1.601*	*-0.106*	—	—	—
First-Stage F-Statistic	163.6	163.6	163.6	163.6	163.6	231.6
Industry FE	yes	yes	yes	yes	yes	yes
Observations	315	315	315	315	315	243
B. Interactions with Leaders / Laggards						
Δln(CHN Imports(-1))						
× Leaders Indicator	-0.904	0.338	-0.314	-0.047	-0.191**	0.181*
	(1.879)	(2.184)	(0.233)	(0.068)	(0.086)	(0.098)
Avg. elasticity (IHS variables)	*-0.904*	*0.338*	*-0.328*	—	—	—
× Laggards Indicator	-3.277***	-2.126**	-0.003	-0.120	-0.245***	0.007
	(0.952)	(0.766)	(0.114)	(0.071)	(0.063)	(0.127)
Avg. elasticity (IHS variables)	*-3.277*	*2.126*	*-0.006*	—	—	—
First-Stage F-Statistic	108.4	108.4	108.4	108.4	108.4	30.2
Industry FE	yes	yes	yes	yes	yes	yes
Observations	315	315	315	315	315	243

Notes: The table compares the baseline innovation results from table 3 with a specification estimated in differences over a three-year windows before and after China joined the WTO in 2001. Industry leaders correspond to the top 10 percent of plants with the highest average TFPQ before 2001. All regressions are run at the plant level, control for the change in the logarithm of employment, and include industry (at the 2-digit level) fixed-effects. Each column shows 2SLS coefficients using (lagged) predicted LASSO imports as an instrument for (lagged) Chinese imports. The (cluster-robust) Kleibergen-Paap rK Wald F-statistic is at the bottom of each column. The corresponding Stock-Yogo value for 10% (15%) maximal IV bias is 16.4 (8.96). Overall innovative spending (column 1), R&D spending (column 2), and patents stock (column 3) are differences of average transformed values using the inverse hyperbolic sine function (IHS, see Bellemare and Wichman, 2020) to account for zeros. Product and process innovation (columns 4-5) are categorical variables taking the value one if the establishment reports successful innovation. Section 3.2 explains the procedure followed to derive the product quality measure. All regressions cluster standard errors at the industry level. Key: ** significant at 1%; ** 5%; * 10%.