Testing the Production Approach to Markup Estimation*

Devesh Raval

Federal Trade Commission

devesh.raval@gmail.com

September 7, 2022

Abstract

Under the production approach to markup estimation, any flexible input should recover the markup. I test this implication using manufacturing datasets from Chile, Colombia, India, Indonesia, the US, and Southern Europe, as well as store-level data from a major US retailer, and overwhelmingly reject that markups estimated using labor and materials have the same distribution. For every dataset, markups estimated using labor are negatively correlated with markups estimated using materials, exhibit greater dispersion, and have opposite time trends. I continue to find stark differences in markups estimated using energy and non-energy raw materials. Non-neutral productivity differences across firms can explain these findings.

*First Version: November 2018. I would like to thank Chris Adams, John Asker, Emek Basker, Ben Bridgman, Allan Collard-Wexler, Emin Dinlersoz, Liran Einav, Amit Gandhi, John Haltiwanger, Berthold Herrendorf, Dan Hosken, Liran Einav, Ethan Kaplan, Rob Kulick, Fernando Luco, Ildiko Magyari, Ryne Marksteiner, Aviv Nevo, Ezra Oberfield, Ariel Pakes, Ted Rosenbaum, Pierre-Daniel Sarte, Dave Schmidt, Marshall Steinbaum, Philippe Sulger, Andrew Sweeting, Nico Trachter, James Traina, Brett Wendling, Kirk White, Nate Wilson, and Mo Xiao, as well as editor Thomas Chaney and three anonyomous referees for their comments on this paper. I also thank Jordi Jaumandreu, Paul Scott, and Dan Ackerberg for discussing this paper at the 2019 IIOC, 2019 NYC IO Day, and 2020 AEA conferences, and Ana Fernandes, Joep Konings, and Benni Moll for providing deflators for various datasets in this paper. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the Federal Trade Commission, or its Commissioners.

Measuring the markup of price over cost is central to recent debates on whether market power has been rising for the US and the world economy (Basu, 2019; Berry et al., 2019; De Loecker et al., 2020; Syverson, 2019). Markups are crucial to evaluate the effects of mergers and changes in trade barriers. Rising aggregate markups can also explain macroeconomic phenomena such as the decline in the labor share of income (Grossman and Oberfield, 2022).

The *production approach* to markup estimation (Hall, 1988; Klette, 1999; De Loecker and Warzynski, 2012) has allowed economists to measure aggregate markups by estimating firm level markups across industries. The production approach uses flexible input choice to identify the markup as a variable input's output elasticity divided by its share of revenue.[1]

This approach requires one to know the production function. In practice, economists using the approach have typically assumed that productivity is Hicks neutral. However, when productivity is labor augmenting, more productive firms will have different output elasticities of labor and materials than less productive firms. Ignoring such heterogeneity will lead to systematically different markups estimated using different inputs.

Because *any* flexible input identifies the markup, the markup is overidentified with multiple flexible inputs. I thus compare markups estimated using labor, materials, or, mirroring cost of goods sold in De Loecker et al. (2020), a composite of both.[2] I conduct these comparisons using manufacturing censuses from Chile, Colombia, India, and Indonesia, firm level

[1] Given competitive input markets, a cost minimizing firm sets the additional revenue from a marginal increase in a flexible input equal to the marginal cost of the input multiplied by the markup.

[2] In the literature, De Loecker and Warzynski (2012) and Blonigen and Pierce (2016) use labor, De Loecker et al. (2016) materials, De Loecker and Scott (2017) both, De Loecker and Eeckhout (2018) cost of goods sold, and De Loecker et al. (2020) cost of goods sold (Compustat) and labor (Economic Census).

data from financial statements for the US and Southern Europe, as well as unique data on individual stores from a nationwide US retailer.

Because the production approach requires estimates of the production function, these comparisons jointly test the assumptions of the production approach itself and auxiliary assumptions on production technology. The validity of estimates of markups using the production approach will thus also depend upon the accuracy of auxiliary assumptions such as Hicks neutrality. I follow De Loecker and Warzynski (2012) and estimate production functions using the Ackerberg et al. (2015) control function estimator, which assumes productivity is Hicks neutral.

I strongly reject that different inputs estimate the same markup in all seven datasets. I focus on three major features of the markup distribution. Labor markups are much more dispersed than materials markups. Markup measures using labor and materials are *negatively* correlated in the cross-section. Finally, their time trends are negatively correlated as well.

One explanation for these findings is that frictions such as hiring and firing costs or monopsony power affect the labor, but not materials, static cost minimization condition. However, I continue to find stark differences between markups estimated using energy and non-energy raw materials. In addition, after controlling for local labor markets using the retailer data – which should capture both differences in input prices and monopsony power across stores – labor markups remain highly negatively correlated with materials markups.

My findings of conflicting correlations when estimating markups with different inputs are also robust to several estimation approaches that assume only neutral productivity differences, estimating production functions at the subindustry or product level, and estimating quantity rather than revenue production functions.

Non-neutral technology can explain these findings. Cost minimization implies that the ratio of output elasticities for two inputs must equal the ratio of their costs. With only Hicks neutral productivity, the ratio of output elasticities is a constant for the Cobb-Douglas, and a deterministic function of production function parameters and inputs with no error for the translog, and so cannot match the variation in input costs in the data. In contrast, non-neutral productivity provides an error term for the ratio of output elasticities that saturates the model.

When labor and materials are complements, higher labor augmenting productivity would both lower labor's output elasticity relative to materials' output elasticity and labor costs relative to materials costs. By ignoring such productivity differences when estimating output elasticities, markups based upon alternative inputs would have opposing time trends and negative correlations, as I have found.

This article is most similar to work that examines differences between markup estimates using the production approach. De Loecker et al. (2020), Karabarbounis and Neiman (2018), and Traina (2018) debate how using different inputs from Compustat affects the aggregate trend in US markups, while Bridgman and Herrendorf (2022) examines the same question using the National Accounts. De Loecker and Scott (2017) find similar average markup estimates using the demand approach, as in Berry et al. (1995), to those from the production approach using data on US breweries.

This article is also related to the literature on non-neutral productivity. Raval (2019) and Oberfield and Raval (2021) document growth in labor augmenting productivity and labor augmenting productivity differences for US manufacturing; Doraszelski and Jaumandreu (2018) and Zhang (2019) do the same using Spanish manufacturing and Chinese steel data.

Section 1 lays out the production approach to estimating markups. Section 2 detail the data and control function estimators, while Section 3 tests the production approach. Section 4 argues that labor augmenting technology differences can explain the failure of the tests and discusses new estimators (Demirer, 2020; Doraszelski and Jaumandreu, 2019) accounting for such differences. Section 5 concludes.

1 Production Approach

Hall (1988) introduced the production approach by showing that, under perfect competition, increases in an input should increase output by the input's share of revenue. He identified the markup as the wedge between the two; that is, the ratio of the change in output divided by the change in the input multiplied by the revenue share. Because productivity shocks would also affect output and inputs, Hall (1988) proposed using aggregate instruments for input growth to estimate the markup.[3]

[3] Hall (1988) estimated this model on time series data for several industries, with aggregate instruments

Klette (1999) was the first to extend the Hall (1988) framework to firm-level data, deriving that the firm level markup is the output elasticity of an input divided by its share of revenue for the firm, as in (4) below. However, he argued that this equation was unlikely to hold without error. Instead, Klette (1999) developed a panel data framework to jointly estimate the average markup across firms in the industry using within firm changes in variable inputs, and average returns to scale using within firm changes in fixed inputs.

Unlike Klette (1999), De Loecker and Warzynski (2012) recovered firm level markups by estimating the output elasticities for each firm. The key assumptions in their model were that the firm cost minimizes in each period with respect to a variable input, and that the firm is a price taker in the input market for that input. Below, I follow De Loecker and Warzynski (2012) and derive the estimator for the markup under these assumptions.

A firm produces output with production function $F_{it}(K_{it}, L_{it}, M_{it})$, where K_{it} is capital for firm i and time t, L_{it} is labor, and M_{it} is materials. The firm receives price P_{it} for its output and faces input prices p_{it}^X for input X. A cost minimizing firm sets marginal products equal to factor prices. This implies, for variable input X_{it},

$$P_{it}\frac{\partial F_{it}}{\partial X_{it}} = \frac{P_{it}}{\lambda_{it}} p_{it}^X, \tag{1}$$

where λ_{it} is the firm's marginal cost.[4] The left hand side is the marginal revenue product of

such as military spending, oil prices, and the president's party. Others followed with more disaggregate industry-level data (Domowitz et al., 1988; Basu and Fernald, 1997).

[4]The marginal cost is the Lagrange multiplier on the production function in the cost minimization prob-

increasing input X_{it}. The right hand side is the marginal cost of increasing X_{it} – its price, p_{it}^X – multiplied by the markup $\frac{P_{it}}{\lambda_{it}}$. Thus, the markup is a wedge between the marginal revenue product of an input and the marginal cost of an input.

Converting this expression to elasticity form[5], the output elasticity for input X, β_{it}^X, is equal to the markup μ_{it} multiplied by input X's share of revenue s_{it}^X:

$$\frac{\partial F_{it}}{\partial X_{it}} \frac{X_{it}}{F_{it}} = \frac{P_{it}}{\lambda_{it}} \frac{p_{it}^X X_{it}}{P_{it} F_{it}} \tag{2}$$

$$\beta_{it}^X = \mu_{it} s_{it}^X. \tag{3}$$

The markup μ_{it} is then the output elasticity of input X divided by X's share of revenue:

$$\mu_{it} = \frac{\beta_{it}^X}{s_{it}^X}. \tag{4}$$

This expression for markups holds for *all* variable inputs at the firm level. Thus, I can test the production approach by examining whether the markup recovered using one input is the same as the markup recovered using another.

lem.

[5]Formally, multiply each side by $\frac{X_{it}}{F_{it}}$ and divide each side by the price P_{it}.

2 Data and Estimation

I summarize the seven datasets used in this paper in Table I, and provide more details on data construction in Appendix A.

The first six datasets cover manufacturing; I use yearly plant censuses for Chile, Colombia, and India, firm census for Indonesia, data from Compustat for an unbalanced panel of US public firms, and data from ORBIS for a balanced panel of all firms in Italy, Spain, and Portugal ("Southern Europe"). I define industries at an equivalent to 2 digit SIC (a broader 2 digit NAICS for the US) and include only industries with at least 1,000 observations.

I also use unique store level data on revenue and inputs for thousands of retail stores from a major nationwide US retailer ("Retailer") for three years. It provides each store's location, which allows me to assess explanations for differences in markups estimated using different inputs in Section 4.2. Because the retailer's data is from its internal accounting systems, it is much higher quality than manufacturing survey data.

Table I Datasets

Dataset	Unit of Observation	Time Period	No. Establishments	No. Industries Used
Chile	Manufacturing Plant	1979-1996	5,000 / year	16
Colombia	Manufacturing Plant	1978-1991	7,000 / year	21
India	Manufacturing Plant	1998-2014	30,000 / year	23
Indonesia	Manufacturing Firm	1991-2000	14,000 / year	22
US (Compustat)	Manufacturing Firm	1970-2010	500 / year	3
Southern Europe (ORBIS)	Manufacturing Firm	2011-2020	100,000 / year	23
Retailer	Retail Store	3 years	Thousands / year	1

I follow De Loecker and Warzynski (2012) and estimate industry-level revenue production functions using the Ackerberg et al. (2015) control function estimator, which imposes that productivity is Hicks neutral and evolves following a Markov process. Appendix B provides more details on estimation.

I then estimate Cobb-Douglas and translog production functions, with two input specifications for each production function. In one specification, inputs are capital, labor, and materials; in another, inputs are capital and a composite variable input of labor and materials costs. I examine a composite variable input in order to mimic cost of goods sold, which combines labor and materials, in De Loecker et al. (2020).[6] With the resulting output elasticities, I build six markup estimates for each establishment-year, using one of three inputs (labor, materials, or the composite input) and one of the two production functions.

3 Empirical Tests

Under the production approach, any flexible input identifies the markup. I test the production approach by examining how markup dispersion, time series correlations, and cross-sectional correlations vary using different inputs. For all of these tests, and in all the datasets, I strongly reject that different inputs estimate the same markup.

In Appendix D, I continue to find sharp differences between inputs examining markup correlations with size, exports, a profit share based markup, and the level of competition.[7]

[6] In some cases, datasets record aggregated variable expenditures of the firm but do not differentiate between inputs like labor and materials.

[7] I also examine average markups in Appendix E.3.

3.1 Dispersion in Markup Estimates

Under the production approach, the degree of markup dispersion should be the same using different flexible inputs. Instead, I find very different levels of dispersion using different inputs. As an illustrative example, I plot the distribution of each markup across manufacturing plants in the Chilean Food Products industry in 1996 using the translog estimates in Figure 1. The red solid lines are the labor markup, the blue dashed lines the materials markup, and the green dash-dot lines the composite variable input markup. The labor markups are much more dispersed than the materials markups, which are in turn more dispersed than the composite input markups. Because of its greater dispersion, a large fraction of labor markups are below one, which might be considered a lower bound on markups (Flynn et al., 2019).

I find a similar pattern in all the datasets. I measure dispersion by calculating the 90/50 ratio of the markup estimates, which I report in Table II.[8] Just as in Figure 1, labor markups are more disperse than materials markups, which are more disperse than composite input markups, for each dataset and production function except for the US Cobb-Douglas estimates. For example, using the translog estimates, the 90th percentile markup is 103% higher than the median markup for Chile using labor, 39% higher using materials, and 17% using the composite input.

For the retailer, there is hardly any dispersion in materials markups – the 90th percentile

[8] I report the 75/25 and 90/10 ratios in Appendix E.2.

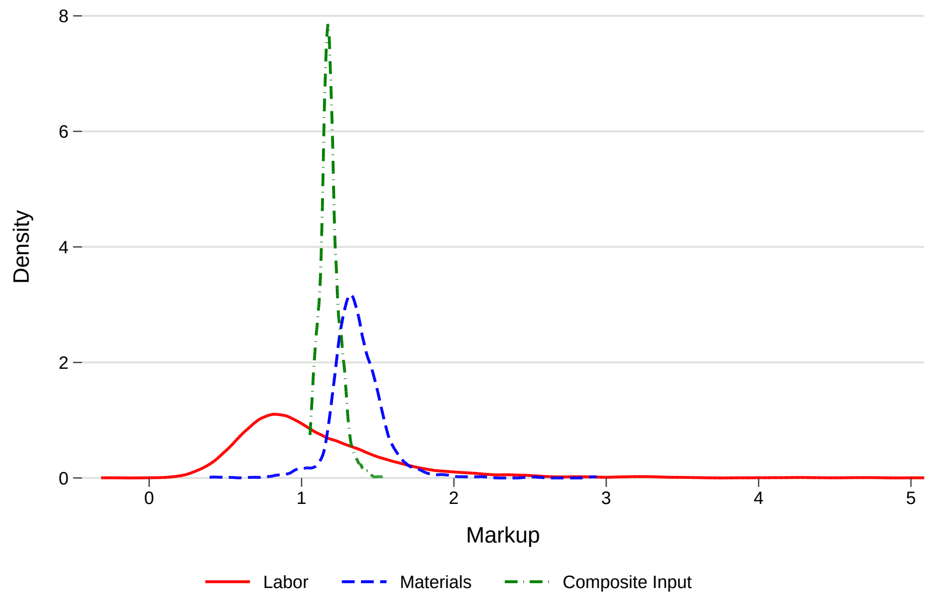


Figure 1 Distribution of Translog Markups for Chilean Food Products, 1996

markup is only 3% higher than the median and 6% higher than the 10th percentile – but substantial dispersion in the labor markup. For the labor markup, the 90th percentile is 30% higher than the median markup and 76% higher than the 10th percentile under the translog estimates.

3.2 Time Trends

Under the production approach, the time path in markups should be the same using different flexible inputs. To test this, I estimate the following specification:

$$\log(\mu_{it}^X) = \alpha + \gamma_t + \delta_n + \epsilon_{it} \tag{5}$$

Table II 90/50 Ratio of Markup Estimates

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	2.67	2.03	1.53	1.39	1.17	1.17
	(0.013)	(0.008)	(0.003)	(0.004)	(0.001)	(0.001)
Colombia	2.88	1.82	1.82	1.43	1.16	1.17
	(0.016)	(0.005)	(0.008)	(0.004)	(0.001)	(0.001)
India	4.04	2.95	1.38	1.29	1.14	1.14
	(0.013)	(0.007)	(0.001)	(0.001)	(0.000)	(0.000)
Indonesia	4.06	3.12	1.66	1.46	1.15	1.16
	(0.025)	(0.019)	(0.004)	(0.003)	(0.001)	(0.001)
US	2.30	3.44	2.61	2.25	1.26	2.13
	(0.017)	(0.040)	(0.026)	(0.022)	(0.003)	(0.009)
S Europe	2.44	1.79	1.83	1.27	1.11	1.10
	(0.002)	(0.001)	(0.002)	(0.001)	(0.000)	(0.000)
Retailer	1.23	1.30	1.02	1.03	1.02	1.02
	(0.002)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Estimates use all establishments and years. Standard errors are based on 20 bootstrap simulations. For India, these estimates ignore the sample weights.

where μ_{it}^X is the markup using input X for establishment i in year t, and γ_t and δ_n are year and industry fixed effects. I then plot the year effects using the translog estimates in Figure 2, with the first year normalized to zero. The red solid lines are the labor markup, the blue dashed lines the materials markup, and the green dash-dot lines the composite input markups.[9]

For all of the datasets, I find *opposing* patterns over time using labor compared to materials to measure the markup. The time trend for composite input markups lie between the two, but much closer to materials, and exhibit less extreme movements.

For example, for Colombia, the average labor markup falls by 28% over the sample,

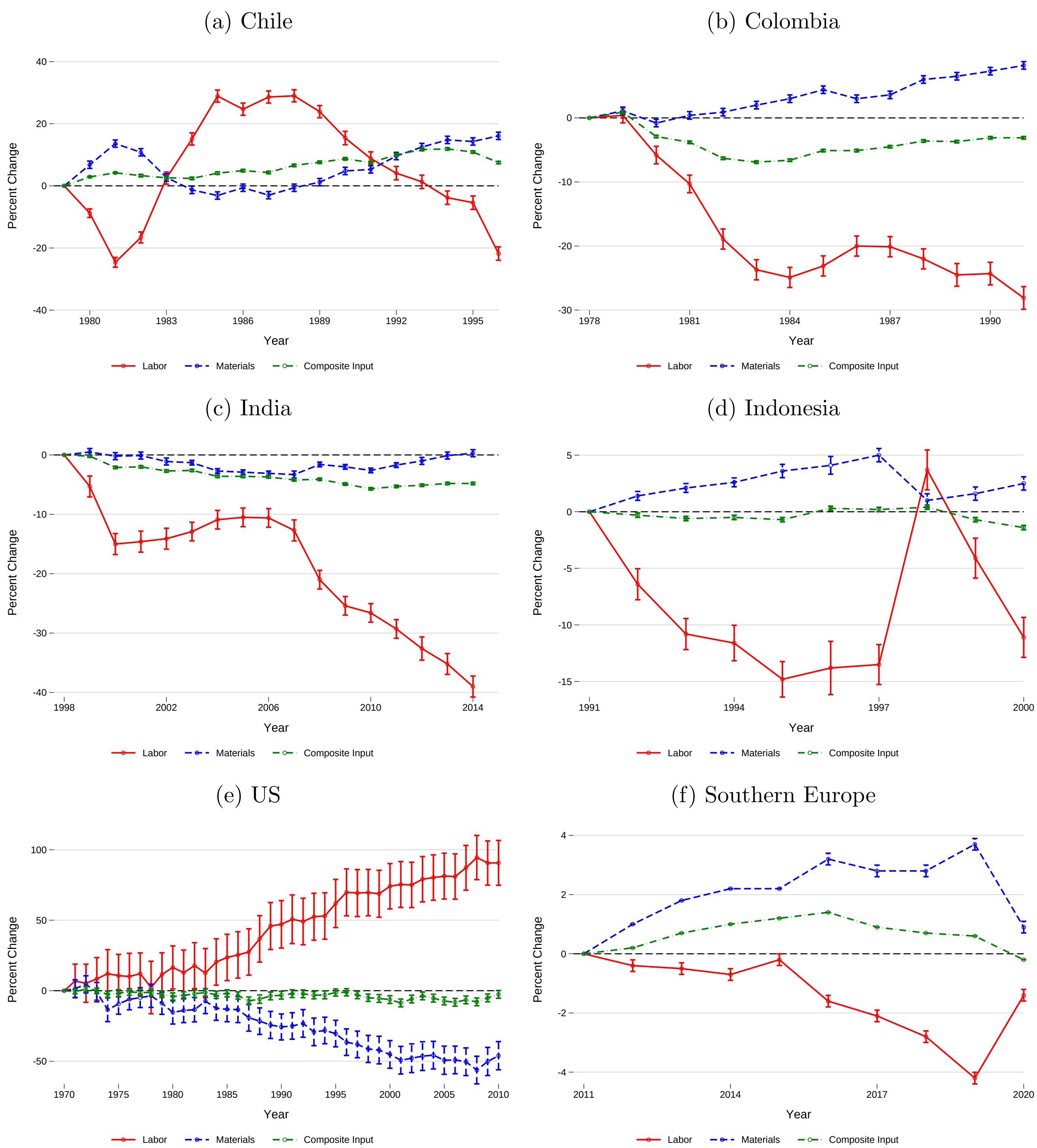
[9] I include the Cobb-Douglas trends in Figure 10 in Appendix E.1. I always find significantly different markup trends using different inputs.

while the average materials markup rises by 8% and the composite input markup declines by 3%. For India, the average labor markup is 39% lower at the end of the sample, while the materials markup exhibits little change and the composite input markup declines by 5%. For Indonesia, the Asian financial crisis strikes in 1998. The average labor markup rises by 17% in 1998, while the average materials markup declines by 4% and the composite input markup remains unchanges. Except for Southern Europe, estimated markups exhibit large changes over the sample period.

A major advantage of the production approach to markup estimation has been the ability to aggregate across producers to estimate the aggregate markup, which is useful to answer many macroeconomic questions, such as reasons for the decline in the labor share. While De Loecker et al. (2020) aggregate by weighting markups by sales, Edmond et al. (2018) argue that the cost weighted markup is the right benchmark for the aggregate markup for welfare calculations of the cost of market power.

I compare changes in the unweighted aggregate markup to a cost aggregated markup in Figure 3. The unweighted average labor markup increases by 91% from 1970 to 2010, compared to a decline of 46% with the materials markup and 3% for the composite input markup. With cost weights, the markup increase is larger; the aggregate labor markup rises by 154%, while the materials markup declines by 23% and composite input markup rises by 3%. In Appendix E.4, I show that using different inputs to estimate markups continues to lead to very different trends after aggregating with cost weights, or with sales weights, for

Figure 2 Markup Time Trends using Translog Estimates

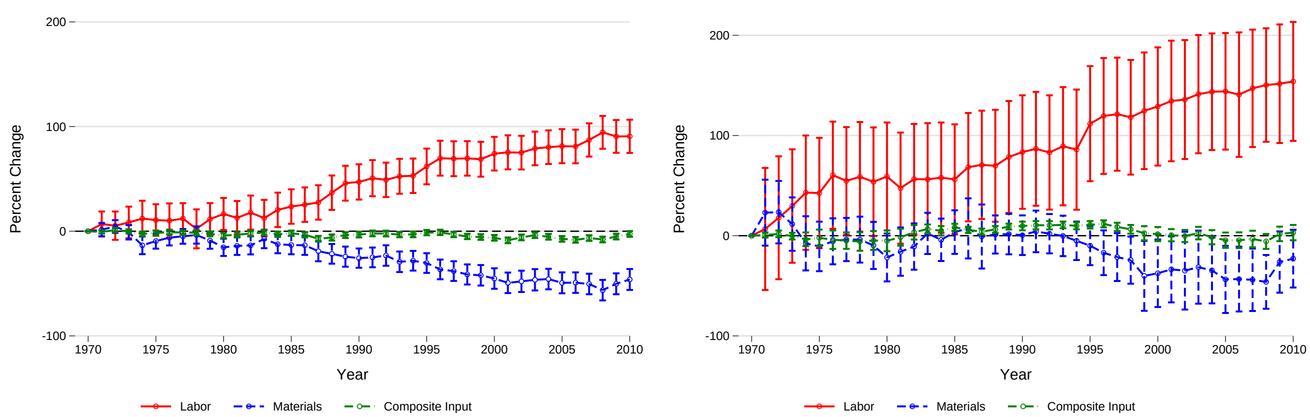


Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

all of the datasets.

Figure 3 Markup Time Trends using Translog Estimates: US

(a) Unweighted (b) Cost-Weighted



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero. The left figure weights all firms equally, whereas the right figure weights each firm by its share of cost.

3.3 Correlations of Markup Estimates

Under the production approach, markup estimates using different inputs for the same establishment should be highly correlated with each other. Instead, I find negative correlations between labor and materials markups. For example, in Figure 4, I plot the materials markup on the x-axis against the labor markup on the y-axis for all plants in the Chilean Food Products industry in 1996 using the translog estimates. Each a point is a different manufacturing plant with the best linear fit as a solid black line. There is a slight negative relationship between the labor markup and materials markup.

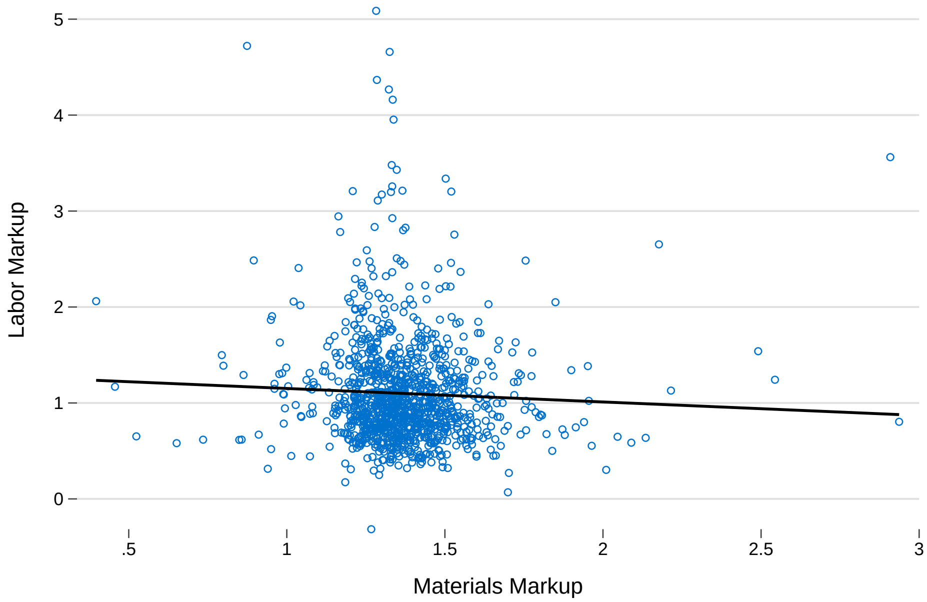


Figure 4 Correlation of Markups for Chilean Food Products, 1996

Note: Each point is the translog markup for a manufacturing plant in Chilean Food Products in 1996; the x-axis is the materials markup and the y-axis is the labor markup. Solid black line is the the best linear fit.

I examine the correlation between markup estimates for all the datasets by estimating the following regression:

$$\log(\mu_{it}^{L}) = \alpha + \beta \log(\mu_{it}^{M}) + \gamma_t + \delta_n + \epsilon_{it} \tag{6}$$

where μ_{it}^{L} and μ_{it}^{M} are the markups using labor and materials for establishment i in year t. I also include controls γ_t and δ_n, year and industry fixed effects, so estimated correlations do not reflect the yearly trends discussed in the previous section. In this specification, β represents the elasticity of the markup using labor with respect to the markup using materials.

I report these correlations between markup measures in Table III. The labor and materials markups are *negatively* correlated with each other, the opposite of the relationship implied by the production approach. Under the translog estimates, an establishment with a 1% higher materials markup has, on average, a 0.16% lower labor markup for Chile, 0.28% lower for Colombia, 0.53% lower for India, 0.48% lower for Indonesia, 0.50% for the US, 0.32% for Southern Europe, and 10.08% lower for the Retailer. In general, the magnitude of the negative correlation is even higher using the Cobb-Douglas estimates.[10]

Table III Relationship between Markup Estimates

Dataset	Cobb-Douglas	Translog
Chile	-0.66	-0.16
	(0.017)	(0.014)
Colombia	-0.99	-0.28
	(0.015)	(0.021)
India	-1.73	-0.53
	(0.012)	(0.009)
Indonesia	-0.97	-0.48
	(0.018)	(0.021)
US	-0.20	-0.50
	(0.032)	(0.029)
S Europe	-0.80	-0.32
	(0.005)	(0.008)
Retailer	-7.51	-10.08
	(0.143)	(0.102)

Note: Estimates based on (6) where the labor markup is the dependent variable and materials markup the independent variable. Standard errors are clustered at the establishment level.

[10] The large magnitude of the elasticities for the retailer is due to the measurement error correction to the input share of revenue as in (15), because the estimated measurement error in sales is negatively correlated with the materials share of revenue. If I ignore this correction, the elasticity between the labor and materials markup is -1 for the Cobb-Douglas case and -2.3 for the translog case.

3.4 Energy

One potential explanation for these findings is that labor is not a flexible input. The literature suggests that violations of the static first order conditions are likely to be more severe for labor (Dobbelaere and Mairesse, 2013), either due to hiring and firing costs when adjusting labor (Petrin and Sivadasan, 2013), bargaining with unions, or labor monopsony power.[11]

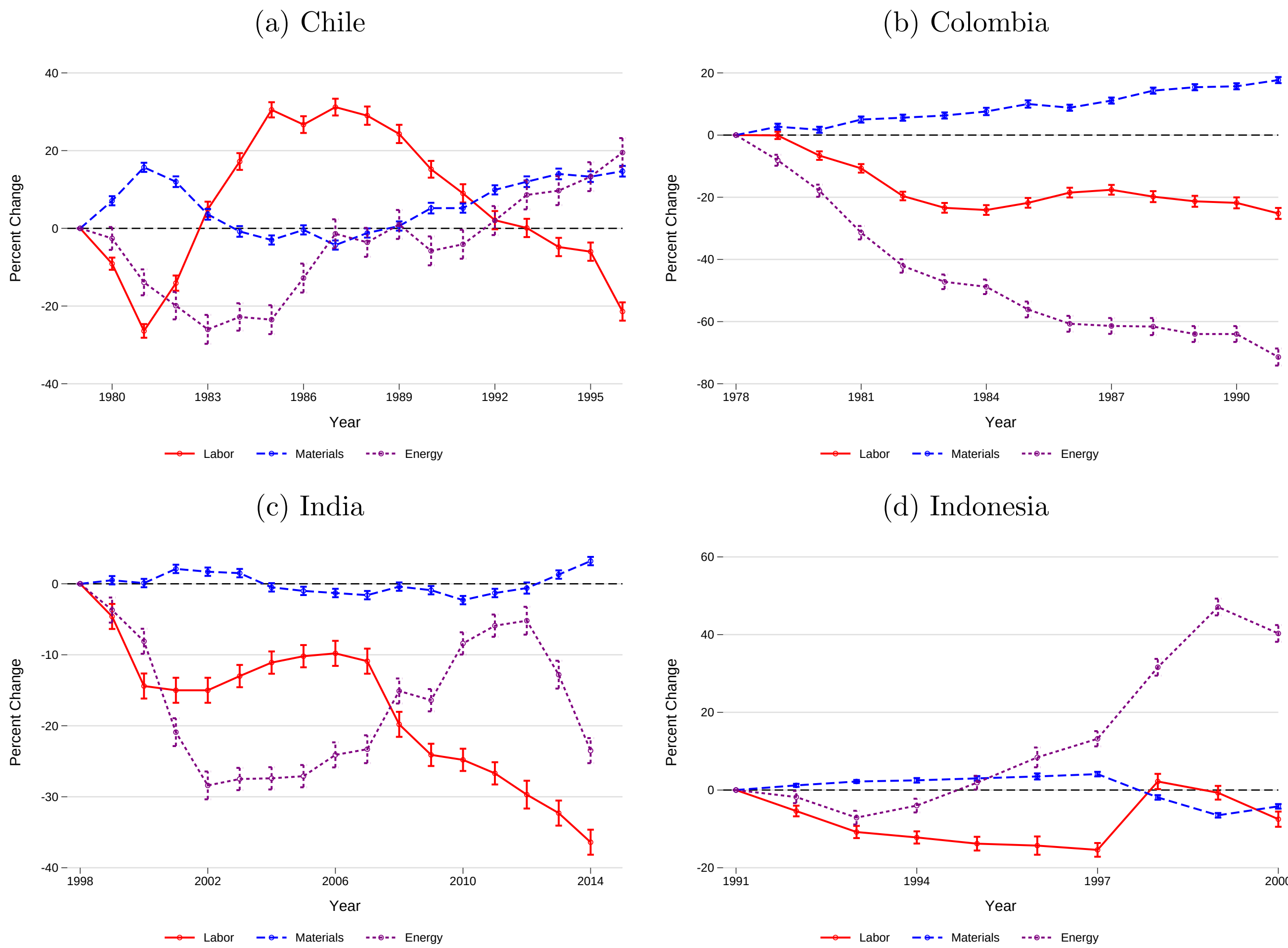
Thus, for the four manufacturing censuses, I separate materials into raw materials and energy, where energy includes both electricity and fuel expenditure, and examine markups for each separately. Both raw materials and energy should be robust to labor-specific violations of the static cost minimization conditions. I then estimate production functions with capital, labor, raw materials, and energy as separate flexible inputs.

I examine time trends separating raw materials and energy estimated using (5). I depict the translog estimates in Figure 5, and the Cobb-Douglas figures in Figure 11. In all four datasets, the raw materials markup has a different time trend than the energy markup.

I report correlations between markup estimates using (6) in Table IV; for example, "Labor on Energy" indicates that the (logged) labor markup is the dependent variable and energy markup the independent variable. Neither the labor or raw materials markup is highly correlated with the energy markup. The raw materials markup is uncorrelated with the energy markup using the translog estimates. The labor markup is negatively correlated with

[11] Union bargaining under a "right to manage" model, in which bargaining is over the wage but the firm can freely choose the number of workers, does not violate my baseline approach. See Nickell and Andrews (1983) and Dobbelaere and Mairesse (2013).

Figure 5 Markup Time Trends using Translog Estimates, with Energy



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

the energy markup under the translog estimates, with a 0% increase in the energy markup leads, on average, to a 0.02% to 0.1% decline in the labor markup. These findings are inconsistent with purely labor-specific violations of the cost minimization conditions.

Table IV Relationship between Markup Estimates: Energy and Raw Materials Separated

	Labor on Raw Materials		Labor on Energy		Raw Materials on Energy	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	-0.60	-0.05	0.21	-0.08	-0.13	-0.01
	(0.017)	(0.013)	(0.008)	(0.006)	(0.003)	(0.002)
Colombia	-0.71	-0.05	0.16	-0.05	-0.26	0.00
	(0.014)	(0.011)	(0.006)	(0.005)	(0.006)	(0.003)
India	-1.38	-0.32	0.28	-0.12	-0.11	0.00
	(0.019)	(0.008)	(0.003)	(0.003)	(0.001)	(0.001)
Indonesia	-0.75	-0.18	0.16	-0.10	-0.14	0.01
	(0.023)	(0.019)	(0.005)	(0.006)	(0.002)	(0.002)

Note: Estimates based on (6) for markups from two flexible inputs, so Labor on Raw Materials indicates a regression where the labor markup is the dependent variable and raw materials markup the independent variable. Standard errors are clustered at the establishment level.

3.5 Robustness

In Appendix C, I show that the large, substantive differences between markups estimated with different inputs demonstrated in this section are robust to several additional specifications. First, these patterns are robust to using several alternative production function estimators assuming neutral productivity, including a dynamic panel approach (Blundell and Bond, 2000), an alternative control function approach (Flynn et al., 2019), and a industry-year level cost share approach. Second, these patterns continue to hold estimating production functions at the subindustry or product level. Third, I find similar patterns estimating quantity rather than revenue production functions using a set of Indian homogenous products. Finally, the data patterns are not consistent with measurement error explanations.

4 Non-Neutral Productivity and Markups

Why are markups estimated using labor so different than those estimated using materials? Dividing the labor markup by the materials markup, we have[12]:

$$\frac{\hat{\mu}_{it}^{L}}{\hat{\mu}_{it}^{M}} = \frac{\beta_{it}^{L}}{\beta_{it}^{M}} \frac{s_{it}^{M}}{s_{it}^{L}} = \frac{\beta_{it}^{L}}{\beta_{it}^{M}} \left(\frac{w_{it} L_{it}}{p_{it}^{m} M_{it}}\right)^{-1}, \tag{7}$$

so the ratio of markups is equal to the ratio of output elasticities divided by the labor cost to materials cost ratio. To estimate the same markup using different inputs, the labor cost to materials cost ratio has to equal the ratio of output elasticities. For a Cobb-Douglas production function, the ratio of output elasticities is a constant $(\frac{\beta_l}{\beta_m})$, so any differences in the labor to materials cost ratio across plants violate this equality and will show up as differences in estimated markups. A translog production function allows the output elasticities to vary based upon inputs, but they remain a deterministic function of production parameters and inputs with no error term. Thus, the translog estimated output elasticities cannot capture the full degree of heterogeneity in input shares across plants.[13]

[12] I would like to thank an anonymous referee for suggesting this framing of the problem.

[13] The ratio of output elasticities for the translog is:

$$\frac{\beta_l + 2\beta_{ll} l_{it} + \beta_{kl} k_{it} + \beta_{ml} m_{it}}{\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}}.$$

So far, I have assumed Hicks neutral productivity, so productivity differences do not affect output elasticities. Below, I show that non-neutral productivity differences can fully capture the heterogeneity in input shares across plants by providing an error term in the ratio of output elasticities.

4.1 Theory

In order to allow productivity to be non-neutral, I assume a CES production function with elasticity of substitution σ, neutral productivity A_{it}, labor augmenting productivity B_{it}, and distribution parameters α_l and α_m:

$$F_{it} = A_{it}((1-\alpha_l-\alpha_m)K_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l(B_{it}L_{it})^{\frac{\sigma-1}{\sigma}} + \alpha_m M_{it}^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}. \tag{8}$$

Input revenue shares are equal to the input's output elasticity divided by the markup μ_{it}:

$$\frac{w_{it}L_{it}}{P_{it}F_{it}} = \frac{1}{\mu_{it}}(\frac{w_{it}}{\lambda_{it}A_{it}})^{1-\sigma}(\alpha_l)^{\sigma}(B_{it})^{\sigma-1} \tag{9}$$

$$\frac{p_{it}^m M_{it}}{P_{it}F_{it}} = \frac{1}{\mu_{it}}(\frac{p_{it}^m}{\lambda_{it}A_{it}})^{1-\sigma}(\alpha_m)^{\sigma} \tag{10}$$

where λ_{it} is the marginal cost, w_{it} the wage, and p_{it}^m the price of materials. An increase in neutral productivity A_{it} does not affect input shares of revenue, as the marginal cost λ_{it} falls to exactly compensate.

Labor augmenting productivity, in contrast, does affect input shares of revenue. An increase in B_{it} is akin to more labor. Thus, after an increase in B_{it}, a firm will increase materials M_{it} to exactly match the increase in effective labor $B_{it}L_{it}$. However, the increase in B_{it} also reduces the cost of an efficient unit of labor, which is $\frac{w_{it}}{B_{it}}$. The plant will then substitute towards relatively cheaper labor, with the ratio of effective labor to materials $\frac{B_{it}L_{it}}{M_{it}}$ changing by σ given the change in the ratio of prices $(w_{it}/B_{it})/p^m_{it}$. Hence the labor cost to materials cost ratio $w_{it}L/p^m_{it}M_{it}$ decreases 1 by a direct effect and increases σ by a substitution effect when B_{it} increases.

When inputs are gross complements, as estimated in Doraszelski and Jaumandreu (2018) and Raval (2019), σ is less than one and so the direct effect is stronger than the substitution effect. A plant with higher labor augmenting productivity will then have a lower labor share, higher materials share, and lower labor cost to materials cost ratio.

Thus, changes in labor augmenting productivity B_{it} move the output elasticities of labor and materials in different directions. In the case when σ is less than one, improvements in B_{it} decrease labor's output elasticity, but increase materials's output elasticity as the marginal cost of production λ falls. If production function estimates ignore labor augmenting productivity differences, a plant with a higher B_{it} would have a lower labor share and higher materials share, and so a higher labor markup and lower materials markup. Estimated markups estimated using different inputs would be negatively correlated.

In Appendix F, I develop a Monte Carlo in which firms have different labor augmenting productivities and set different markups. In the Monte Carlo, control function estimators assuming neutral productivity imply negatively correlated labor and materials markups. In addition, neither labor or materials markups are highly correlated with the true markup.

The revenue share of the composite input of labor and materials is equal to the sum of the right hand side of (9) and (10). Thus, the output elasticity of the combined input also declines as labor augmenting productivity B_{it} increases, albeit less than the labor output elasticity.[14] If not accounted for, labor augmenting productivity differences will also bias the composite input markup. For example, improvements in labor augmenting productivity over time, all other things equal, would decrease the output elasticity of the composite input, and so increase aggregate markups.

4.2 Alternative Explanations

Given (9) and (10), the labor cost to materials cost ratio is:

$$\frac{w_{it}L_{it}}{p^m_{it}M_{it}} = (\frac{w_{it}}{p^m_{it}})^{1-\sigma}(\frac{\alpha_l}{\alpha_m})^{\sigma}(B_{it})^{\sigma-1},$$

so non-neutral productivity differences B_{it} can clearly soak up any differences in the labor cost to materials cost ratio across plants. However, unobserved differences in input prices $\frac{w_{it}}{p^m_{it}}$ could also lead to differences in factor costs (Grieco et al., 2016), as could monopsonistic

[14]The composite input elasticity is $\frac{(w_{it}/B_{it})^{1-\sigma}\alpha_l^{\sigma}+(p^m_{it})^{1-\sigma}\alpha_m^{\sigma}}{(\lambda_{it}A_{it})^{1-\sigma}}$.

behavior by firms that would show up as "wedges" in input first order conditions. Finally, measurement errors in inputs would also lead to differences in factor cost ratios across plants.

In general, it is difficult to assess whether differences in factor cost ratios across plants are due to productivity differences as opposed to these other factors, as we typically do not observe the plant's input prices or its degree of monopsony power. Manufacturing survey data also contain a substantial amount of measurement error (White et al., 2016).

With the retailer, however, I can control for these alternative explanations. First, the retailer's data is based on the internal records of the firm and so should have very little measurement error compared to self-reported survey data. Second, I have detailed data on the store location that allows me to control for local labor markets; the same chain's input prices and monopsony power should not vary within a local labor market. I control for local labor markets through two types of location-year fixed effects. First, I use the MSA of the store, defined as either the Metropolitan Statistical Area or Micropolitan Statistical Area of the retail store's location.[15] Second, I use the store's district as defined by the internal structure of the retailer; each district has about 10 to 20 stores.

I then use the translog estimates and re-estimate (6) after controlling for either MSA-year or district-year fixed effects. A retail store with a 100% higher materials markup has, on average, a 1008% lower labor markup in the previous, baseline estimates, compared to 992% lower with MSA-year fixed effects and 1006% lower with district-year fixed effects. Because

[15] For retail stores not located in a Metropolitan Statistical Area or Micropolitan Statistical Area, the fixed effect is for all non-MSA locations in the same state.

the negative correlation between labor and materials markups remains almost unchanged after controlling for local labor markets, differences in input prices or monoposony power are unlikely to explain the patterns that I find for the retail store.

On the other hand, there are several reasons why stores in the same retail chain may vary in productivity. First, each store is run by its own store manager, and differences in managerial ability are well known to affect productivity (Bloom and Van Reenen, 2007; Fenizia, 2022), such as by affecting staff turnover (Hoffman and Tadelis, 2021). Second, each retail store is a multi-product producer, with each section of the store using its own production function and potentially different types of labor; stores vary in the importance of components of the store.[16] Finally, this retail chain operates some 24-hour stores, which will have a different production technology than stores open during more standard hours.

4.3 Estimation

How can we estimate the production function given non-neutral productivity differences? One possibility is to directly estimate first order conditions such as (9) and (10) using variation in factor prices across firms and then recover the elasticity of substitution and productivities. For example, Raval (2019) exploits cross-sectional variation in wages across US locations together with several instruments for such wages – local amenities and labor demand instruments from Bartik (1991) and Beaudry et al. (2012) – for identification.

[16] This type of heterogeneity is common in retail. For example, Walmart Superstores have a much larger grocery component than standard Walmart stores.

However, the researcher often does not have exogenous variation in input prices, or wants to assume a different production function than the CES. Thus, three new approaches modify existing methods of production function estimation to allow non-neutral productivity.

First, Doraszelski and Jaumandreu (2019) provide a new dynamic panel estimator for markups. The estimator assumes a translog production function in which the combined output elasticity of materials and labor is constant, quantity produced is observable, and neutral productivity follows an autoregressive process. Log quantity produced is then a linear function of its lag and differences in inputs and input shares, and model parameters can be estimated via GMM. Output elasticities are measured with error (similar to Klette (1999)), so individual markups are recovered up to this error.

Second, Demirer (2020) develops a non-parametric control function approach to estimate output elasticities given non-neutral productivity. This approach requires only homothetic separability between labor and materials, together with input choice timing assumptions. As in (7), the ratio of input costs identifies the ratio of output elasticities. Demirer (2020) uses the materials-labor ratio equation to form a control function for labor augmenting productivity, and then the materials demand equation to form a control function for neutral productivity.

Lastly, Raval (2022) generalizes the cost share approach to markup estimation to account for labor augmenting productivity differences. This flexible cost share estimator assumes that cost minimization conditions hold for all inputs on average and returns to scale are constant.

Because labor augmenting productivity is proportional to the ratio of labor to materials costs, Raval (2022) groups plants into bins with similar labor augmenting productivities based upon this ratio and estimates output elasticities from cost shares within each group.

The attractiveness of these approaches will depend upon which assumptions the econometrician is comfortable with placing on the data. For example, Doraszelski and Jaumandreu (2019) and Demirer (2020) estimate an output equation and so require data on quantity produced, whereas Raval (2022) does not require quantity data as he estimates output elasticities only using cost data. Doraszelski and Jaumandreu (2019) and Raval (2022) assume all plants have the same variable returns to scale and total returns to scale, respectively. And Raval (2022) implicitly places strong assumptions on the adjustment process of capital.

5 Conclusion

A key advantage of the production approach to estimating markups is that it allows one to estimate markups across widely differing industries, and thus estimate the aggregate markup. However, the production approach, as currently implemented, delivers very different markups using alternative flexible inputs. Labor markups are negatively correlated with materials markups, have opposing time trends, and are much more disperse. Non-neutral technological differences across plants can explain these findings, as ignoring such differences will lead to negatively correlated labor and materials markups.

The development of the parallel demand approach to markup estimation provides guidance on how to measure markups going forward. The demand approach focuses on modeling the heterogeneity in preferences across consumers; for example, Berry et al. (1995) estimate random coefficients that allow consumers to vary in their sensitivity to price. In order to use the production approach, economists will have to allow more heterogeneity in technology.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Alcott, Hunt, Allan Collard-Wexler, and Stephen O'Connell, "How Do Electricity Shortages Affect Industry? Evidence from India," *American Economic Review*, 2015.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, W. E. Upjohn Institute for Employment Research, 1991.

Basu, Susanto, "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence," *Journal of Economic Perspectives*, August 2019, *33* (3), 3–22.

_ **and John G Fernald**, "Returns to scale in US production: Estimates and implications," *Journal of political economy*, 1997, *105* (2), 249–283.

Beaudry, Paul, David A. Green, and Benjamin Sand, "Does Industrial Composition Matter For Wages? A Test Of Search And Bargaining Theory," *Econometrica*, 2012, *80* (3), 1063–1104.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

_ **, Martin Gaynor, and Fiona Scott Morton**, "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization," *Journal of Economic Perspectives*, August 2019, *33* (3), 44–68.

Blonigen, Bruce A and Justin R Pierce, "Evidence for the Effects of Mergers on Market Power and Efficiency," Technical Report, National Bureau of Economic Research 2016.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The quarterly journal of Economics*, 2007, *122* (4), 1351–1408.

Blundell, Richard and Stephen Bond, "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 2000, *19* (3), 321–340.

Bond, Stephen and Måns Söderbom, "Adjustment Costs and the Identification of Cobb Douglas Production Functions," Technical Report, IFS Working Papers 2005.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch, "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data," Technical Report, NBER Working Paper 2020.

Bridgman, Benjamin and Berthold Herrendorf, "Labor Share, Markups, and Input-Output Linkages: Evidence from the National Accounts," *Mimeo*, 2022.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in US Retail Chains," Technical Report, National Bureau of Economic Research 2017.

Demirer, Mert, "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups," Technical Report, Mimeo 2020.

Diewert, W Erwin and Denis A Lawrence, "Progress in Measuring the Price and Quantity of Capital," *Econometrics*, 2000, *2*, 273–326.

Dobbelaere, Sabien and Jacques Mairesse, "Panel Data Estimates of the Production Function and Product and Labor Market Imperfections," *Journal of Applied Econometrics*, 2013, *28* (1), 1–46.

Domowitz, Ian, R Glenn Hubbard, and Bruce C Petersen, "Market Structure and Cyclical Fluctuations in US Manufacturing," *The Review of Economics and Statistics*, 1988, pp. 55–66.

Doraszelski, Ulrich and Jordi Jaumandreu, "Measuring the Bias of Technological Change," *Journal of Political Economy*, 2018, *126* (3), 1027–1084.

_ and _ , "Using Cost Minimization to Estimate Markups," 2019.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How Costly are Markups?," Technical Report, National Bureau of Economic Research 2018.

Fenizia, Alessandra, "Managers and productivity in the public sector," *Econometrica*, 2022, *90* (3), 1063–1084.

Flynn, Zach, Amit Gandhi, and James Traina, "Measuring Markups with Production Data," 2019.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Foster, Lucia S, Cheryl A Grim, John Haltiwanger, and Zoltan Wolf, "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights," Working Paper 23666, National Bureau of Economic Research 2017.

Gandhi, Amit, Salvador Navarro, and David Rivers, "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, forthcoming.

Greenstreet, David, "Exploiting Sequential Learning to Estimate Establishment-Level Productivity Dynamics and Decision Rules," 2007. Mimeo.

Grieco, Paul LE, Shengyu Li, and Hongsong Zhang, "Production function estimation with unobserved input price dispersion," *International Economic Review*, 2016, *57* (2), 665–690.

Grossman, Gene M. and Ezra Oberfield, "The Elusive Explanation for the Declining Labor Share," *Annual Review of Economics*, 2022, *14* (1), null.

Gutiérrez, Germán and Thomas Philippon, "Investment-less Growth: An Empirical Investigation," Technical Report, National Bureau of Economic Research 2016.

Hall, Robert E, "The Relation Between Price and Marginal Cost in US Industry," *Journal of Political Economy*, 1988, *96* (5), 921–947.

Harper, Michael J., Ernst R. Berndt, and David O. Wood, "Rates of Return and Capital Aggregation Using Alternative Rental Prices," in D.W. Jorgenson and R. London, eds., *Technology and Capital Formation*, Cambridge, MA: MIT Press, 1989.

Hoffman, Mitchell and Steven Tadelis, "People management skills, employee attrition, and manager rewards: An empirical analysis," *Journal of Political Economy*, 2021, *129* (1), 243–285.

Karabarbounis, Loukas and Brent Neiman, "Accounting for Factorless Income," Technical Report, National Bureau of Economic Research 2018.

Keller, Wolfgang and Stephen R Yeaple, "Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States," *The review of economics and statistics*, 2009, *91* (4), 821–831.

Klette, Tor Jakob, "Market power, scale economies and productivity: estimates from a panel of establishment data," *The Journal of Industrial Economics*, 1999, *47* (4), 451–476.

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

Loecker, Jan De and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–71.

— **and Jan Eeckhout**, "Global market power," Technical Report, National Bureau of Economic Research 2018.

— **and Paul T Scott**, "Estimating Market Power: Evidence from the US Brewing Industry," Technical Report, US Census Bureau, Center for Economic Studies 2017.

— **, Jan Eeckhout, and Gabriel Unger**, "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

— **, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik**, "Prices, Markups, and Trade Reform," *Econometrica*, 2016, *84* (2), 445–510.

Melitz, Marc J and Gianmarco IP Ottaviano, "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 2008, *75* (1), 295–316.

Nickell, Stephen and M Andrews, "Unions, Real Wages, and Employment in Britain 1951-1979," *Oxford Economic Papers*, 1983, *35* (0), 183–206.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

Petrin, Amil and Jagadeesh Sivadasan, "Estimating Lost Output from Allocative Inefficiency, with an Application to Chile and Firing Costs," *Review of Economics and Statistics*, 2013, *95* (1), 286–301.

Raval, Devesh, "The Micro Elasticity of Substitution and Non-Neutral Technology," *RAND Journal of Economics*, 2019, *50* (1), 147–167.

— , "A Flexible Cost Share Approach to Markup Estimation," 2022.

Rovigatti, Gabriele and Vincenzo Mollisi, "Theory and Practice of Total-Factor Productivity Estimation: The Control Function Approach using Stata," *The Stata Journal*, 2018, *18* (3), 618–662.

Syverson, Chad, "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, 2004, *86* (2), 534–550.

— , "Macroeconomics and Market Power: Context, Implications, and Open Questions," *Journal of Economic Perspectives*, August 2019, *33* (3), 23–43.

Traina, James, "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," 2018.

White, T Kirk, Jerome P Reiter, and Amil Petrin, "Imputation in US Manufacturing Data and Its Implications for Productivity Dispersion," *Review of Economics and Statistics*, 2016, (0).

Zhang, Hongsong, "Non-Neutral Technology, Firm Heterogeneity, and Labor Demand," *Journal of Development Economics*, 2019.

A Data

For each dataset used in this paper, I construct capital, labor, materials, and sales at the establishment-year level. I provide further details on data construction in Appendix G.

An establishment is a manufacturing plant for the Chilean, Colombian, and Indian data, a firm for the Indonesian, US, and Southern European data, and a retail store for the retailer. India is a census for plants above 100 employees and a sample for smaller plants; I thus use the provided sampling weights.

I use capital, materials, and output deflators in order to construct consistent measures of inputs and outputs over time, and drop any observations with zero or negative capital, labor, materials, sales, or labor costs. I also drop observations in the bottom 1% and top 1% of labor's share of revenue, materials's share of revenue, and the composite variable input share of revenue for each industry to remove outliers.

For labor, I use the number of workers for Chile, Colombia, Indonesia, US, and Southern Europe, the number of manufacturing worker-days for India, and the total number of hours worked for the retailer. Labor costs are the total of salaries and worker benefits.

For materials, I include expenses for raw materials, electricity, and fuels for the manufacturing censuses. For Southern Europe, materials is total materials costs, while for the retailer, materials is the sum of the cost of goods sold across different parts of the store. The composite variable input is the sum of materials and labor costs.[17]

For the US, I do not have separate data fields on labor costs and materials costs. Instead, I follow Keller and Yeaple (2009) and Demirer (2020) and estimate labor costs as total employees multiplied by an industry-level average wage, and materials costs as cost of goods sold and selling, general, and administrative expenses minus depreciation and labor costs.

For capital, where possible, I construct a perpetual inventory measure of capital for each type of capital and rental rates of capital based on an average real interest rate over time plus depreciation for that type of capital. My measure of capital is then the sum of capital stocks times their rental rates, plus any rental payments for capital.[18]

For the manufacturing datasets, I estimate production functions at the industry level, defined at a similar level to two digit US SIC (i.e., Chilean Food Products).[19] Given the smaller size of the US data, I follow De Loecker et al. (2020) and estimate production functions at the 2 digit NAICS

[17] I deflate this input using the output deflator to match De Loecker et al. (2020)'s treatment of cost of goods sold.

[18] This provides an approximation to a Divisia index for capital given different types of capital. See Diewert and Lawrence (2000) and Harper et al. (1989) for details on capital rental rates and aggregation. For the US, I do not have data on separate types of capital, and use rental rates from Oberfield and Raval (2021). For Southern Europe, I only have book values of capital. For the retailer, I use BLS rental rates for retail trade. See Appendix G for more details on capital construction.

[19] For Chile, Colombia, and Indonesia this is at the three digit ISIC (Rev.2) level, for India at the two digit NIC 08 level, and for Southern Europe at the 2 digit NACE level. Estimating production functions at this level of aggregation is consistent with the production function literature, such as Levinsohn and Petrin (2003) or Gandhi et al. (forthcoming).

level, so there are three industries. I only include industries with at least 1,000 observations over the entire dataset. For the retailer, I estimate a single production function across all retail outlets.

B Estimation

Given (4), estimating the markup requires the input share of revenue and the output elasticity of that input. The input share of revenue, defined as costs for input X divided by total firm revenue, is observed. However, the production function has to be estimated to recover output elasticities. I describe below how De Loecker and Warzynski (2012), and subsequent papers using the production approach such as De Loecker et al. (2020), address this estimation challenge using a control function approach that assumes productivity is Hicks neutral.

B.1 Production Functions

All lower case variables are in logged form, so k_{it} is capital, l_{it} labor, and m_{it} materials. For the Cobb Douglas production function with labor and materials, the (logged) production function excluding the Hicks neutral productivity term is:

$$f_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it}$$

and so the output elasticity for input X is simply β_X. For the translog production function, the production function is:

$$f_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it}$$

and so the output elasticity for each input will depend upon the level of all inputs. For example, the firm's output elasticity for materials would be $\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}$.

For both the Cobb-Douglas and translog production functions, the production function coefficients are not time-varying. However, for the translog, output elasticities can vary over time due to changes in factors.

B.2 Control Function Estimation

I follow De Loecker and Warzynski (2012) and use the Ackerberg et al. (2015) (ACF) estimator for my baseline estimates. The ACF estimator imposes substantial additional assumptions on productivity, including that productivity is Hicks neutral and evolves following a Markov process. In addition, it requires a set of timing assumptions where at least one input is decided at the time the firm learns its productivity shock. I discuss problems with this estimator, and alternative estimation approaches using neutral productivity, in Appendix C.1.

The control function approach assumes that observed revenue includes additive measurement error ϵ_{it}. Thus, given log productivity ω_{it}, measured log revenue y_{it} is:

$$y_{it} = f(k_{it}, l_{it}, m_{it}) + \omega_{it} + \epsilon_{it}. \tag{11}$$

Let materials be the flexible input decided at the time the firm learns its productivity shock. Materials is then a function of the observed inputs and productivity $m_{it} = g(k_{it}, l_{it}, \omega_{it})$. It can then be inverted for productivity, so $\omega_{it} = g^{-1}(k_{it}, l_{it}, m_{it})$.[20]

The first stage of the ACF estimator controls for a flexible form of the inputs to recover the additive measurement error ϵ_{it}. Formally, measured log revenue y_{it} is:

$$y_{it} = f(k_{it}, l_{it}, m_{it}) + g^{-1}(k_{it}, l_{it}, m_{it}) + \epsilon_{it} = h(k_{it}, l_{it}, m_{it}) + \epsilon_{it} \tag{12}$$

Since both the production function and productivity are functions of the inputs, they cannot be separated in the first stage. Instead, the nonparametric function h includes both productivity ω_{it} and the production function f. The measurement error in sales ϵ_{it} is a residual in the first stage equation after controlling for h.[21]

The second major assumption of the ACF approach is that productivity follows a first order Markov process. In my implementation, I further assume an AR(1) process. Formally,

$$\omega_{it} = \rho\omega_{it-1} + \nu_{it} \tag{13}$$

with AR(1) coefficient ρ and productivity innovation ν_{it}. In that case, given knowledge of the production function coefficients β, one can recover the innovation in productivity ν_{it} as:

$$\nu_{it}(\beta) = \omega_{it} - \rho\omega_{it-1} \tag{14}$$

The innovation in productivity is a function of production coefficients β because $\omega_{it} = y_{it} - \epsilon_{it} - f_{it}(\beta)$, and ϵ_{it} was recovered in the first stage.

Because the innovation in productivity is, by construction, independent of inputs chosen before time t, moments of the innovations multiplied by inputs chosen before the productivity innovation, such as $E(\nu_{it}l_{it-1})$ or $E(\nu_{it}k_{it})$, identify the production function coefficients.

For the Cobb-Douglas production function, I use capital and the first lag of materials and labor as instruments. For the translog, I use capital and the first lag of materials and labor, as well as their interactions, as instruments.[22]

Finally, I follow De Loecker and Warzynski (2012) and correct the value of sales in the input share of revenue for the measurement error estimated in the first stage. Thus, for input X, the estimate of the markup is:

$$\hat{\mu_{it}} = \frac{\hat{\beta_i^X}}{s_{it}^X exp(\hat{\epsilon}_{it})}. \tag{15}$$

[20]The g() function can include other determinants of materials as well, such as materials prices.

[21]In practice, I use a third order polynomial in inputs for the function h, and also control for year effects.

[22]For the specification with the composite variable input instead of labor and materials separately, I use the lag of the composite input and its interactions as instruments, symmetrically to the case above.

C Robustness to Empirical Tests (For Online Publication)

In this section, I show that the large, substantive differences between markups estimated with different inputs demonstrated in Section 3 are robust to several additional specifications. First, in Appendix C.1, I show that these patterns hold estimating production functions through several different estimation approach compared to the ACF approach in the main text. Second, in Appendix C.2, I show similar patterns estimating production functions at the subindustry or product level. Third, in Appendix C.3, I show similar patterns estimating quantity as opposed to revenue production functions using data on Indian homogeneous products. Finally, in Appendix C.4, I argue that measurement error is unlikely to explain the patterns that I find.

C.1 Alternative Production Function Estimators

Following De Loecker and Warzynski (2012), I used the control function approach of Ackerberg et al. (2015) to estimate production functions. One explanation for my findings is this estimation approach is misspecified, which could happen for several reasons.

First, the auxiliary assumptions required for the control function approach, such as a Markov assumption on productivity together with timing assumptions on when the firm determines its level of inputs, may not hold. Second, Gandhi et al. (forthcoming) show that the ACF procedure is not identified when applied to gross-output production functions.[23] Third, Flynn et al. (2019), Doraszelski and Jaumandreu (2019), and Bond et al. (2020) show how the ACF procedure can fail to identify production function parameters with non-competitive output markets when the dependent variable is revenue and not quantity produced. Fourth, Rovigatti and Mollisi (2018) find that ACF estimates are quite sensitive to the initial conditions used for optimization. Empirically, Foster et al. (2017) show that estimated output elasticities can vary substantially across different estimation approaches.

To examine whether such issues explain my findings, I examine three additional approaches to production function estimation. First, I use a dynamic panel approach to estimation following Blundell and Bond (2000). Second, Flynn et al. (2019) develop a new method to estimate production functions using a similar set of auxiliary assumptions as Ackerberg et al. (2015) together with constant returns to scale. I use this new method to estimate translog production functions.[24] Finally, I use the cost share approach assuming that productivity differences are neutral using industry-year cost shares, as in De Loecker et al. (2020). The cost share estimates allow the output elasticities of the industry-level production function to change over time, but do not allow

[23] See Bond and Söderbom (2005) for an early critique in this vein. Ackerberg et al. (2015) state that "we would not suggest applying our procedure to gross output production functions that are not Leontief in the intermediate inputs".

[24] This approach does not converge for one industry for Chile, Colombia, and Indonesia, and two industries for India for the labor and materials specification, as well as one industry for Indonesia, seven industries for India, and two industries for Southern Europe in the composite variable input specification.

non-neutral technological differences across establishments in the same year.

Using all three methods, the time trends using different inputs estimated using (5) are very different for all cases except for cost shares for Colombia. I depict these in Figure 6 through Figure 8. In addition, after controlling for time trends, I show in Table V that the labor markup remains negatively correlated with the materials markup, with a decline in the labor markup with a 1% increase in the materials markup ranging from -0.20% to -1% using the dynamic panel approach, -0.17% to -7.05% using the Flynn et al. (2019) approach, and from -0.21% to -1% for the cost share approach.

Thus, alternative production function estimators assuming neutral productivity differences cannot explain the differing markup estimates across variable inputs that I document.

Table V Relationship between Markup Estimates: Alternative Estimators

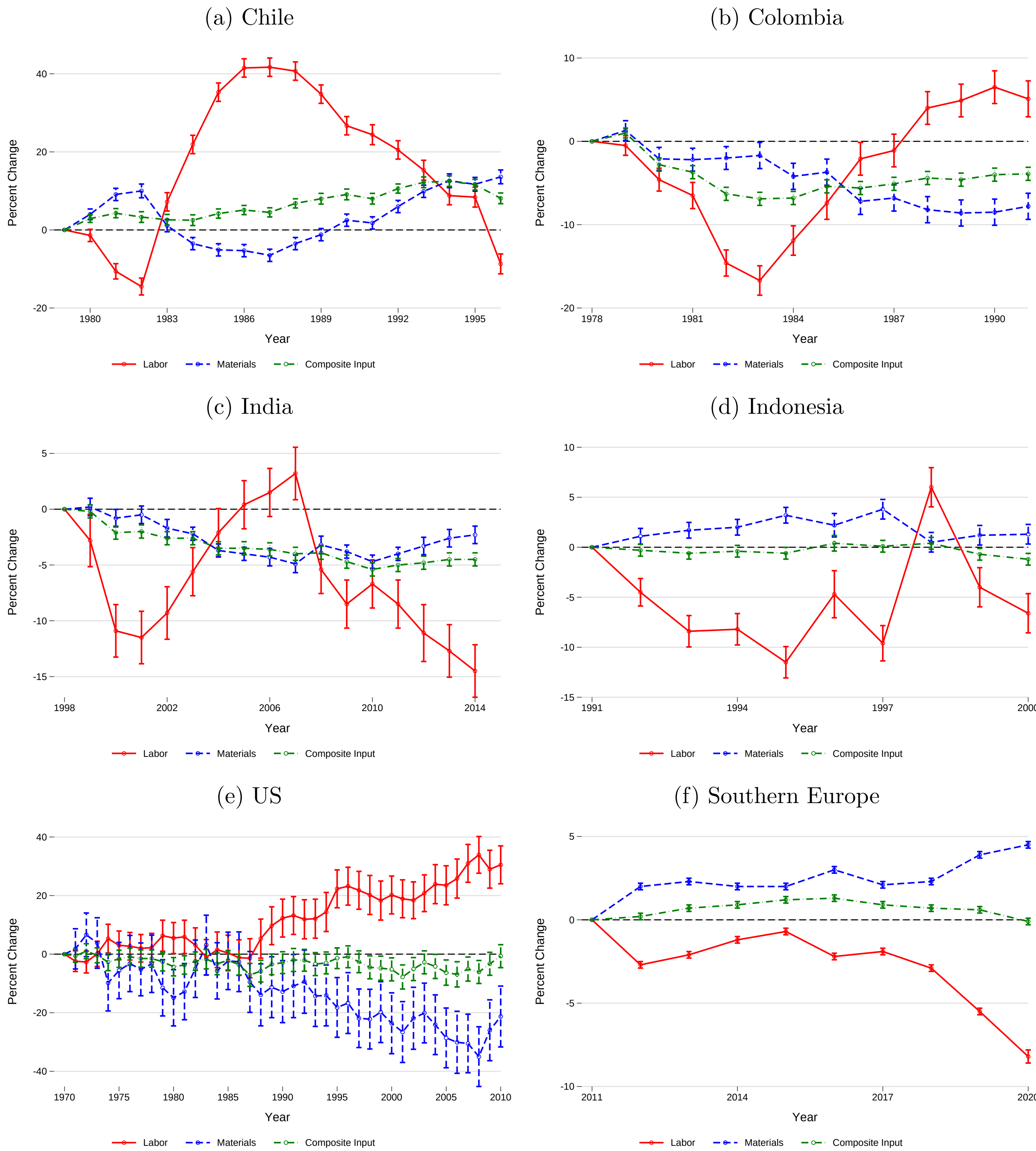
Dataset	Dynamic Panel	FGT	CostShare Ind	CostShare SubInd
Chile	-0.25	-0.69	-0.24	-0.20
	(0.015)	(0.018)	(0.015)	(0.014)
Colombia	-0.65	-1.06	-0.65	-0.61
	(0.008)	(0.020)	(0.008)	(0.009)
India	-0.89	-0.17	-0.89	-0.66
	(0.008)	(0.007)	(0.008)	(0.008)
Indonesia	-0.70	-0.82	-0.51	0.02
	(0.011)	(0.020)	(0.010)	(0.016)
US	-0.20	-0.18	-0.21	-0.21
	(0.021)	(0.032)	(0.021)	(0.022)
S Europe	-0.55	-0.83	-0.55	-0.50
	(0.003)	(0.006)	(0.003)	(0.003)
Retailer	-1.00	-7.05	-1.00	-1.00
	(0.055)	(0.151)	(0.055)	(0.055)

Note: Estimates based on (6) where the labor markup is the dependent variable and materials markup the independent variable. Columns labeled Dynamic Panel are markups based on Blundell and Bond (2000), and labeled FGT based on Flynn et al. (2019), as described in the text. Columns labeled CostShare Ind are markups based on industry-year level cost shares, and CostShare SubInd are markups based on subindustry-year level cost shares, as described in the text. Standard errors are clustered at the establishment level.

C.2 Within Industry Heterogeneity

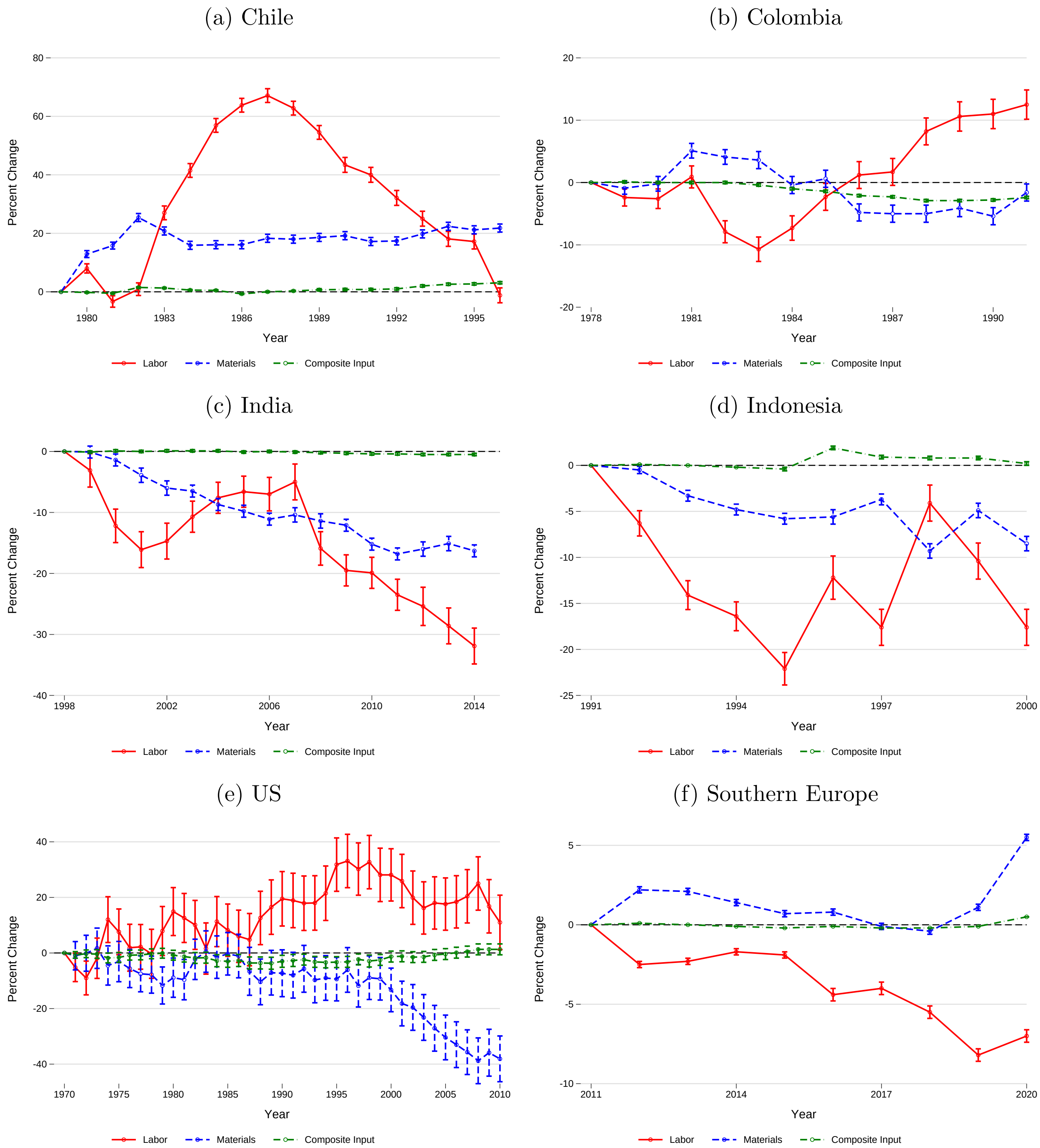
One potential concern is that production functions vary across subindustries or products within a broader industry. With such variation, production function estimates at the industry level may not identify a plant's production function parameters.

Figure 6 Markup Time Trends using Dynamic Panel Estimates



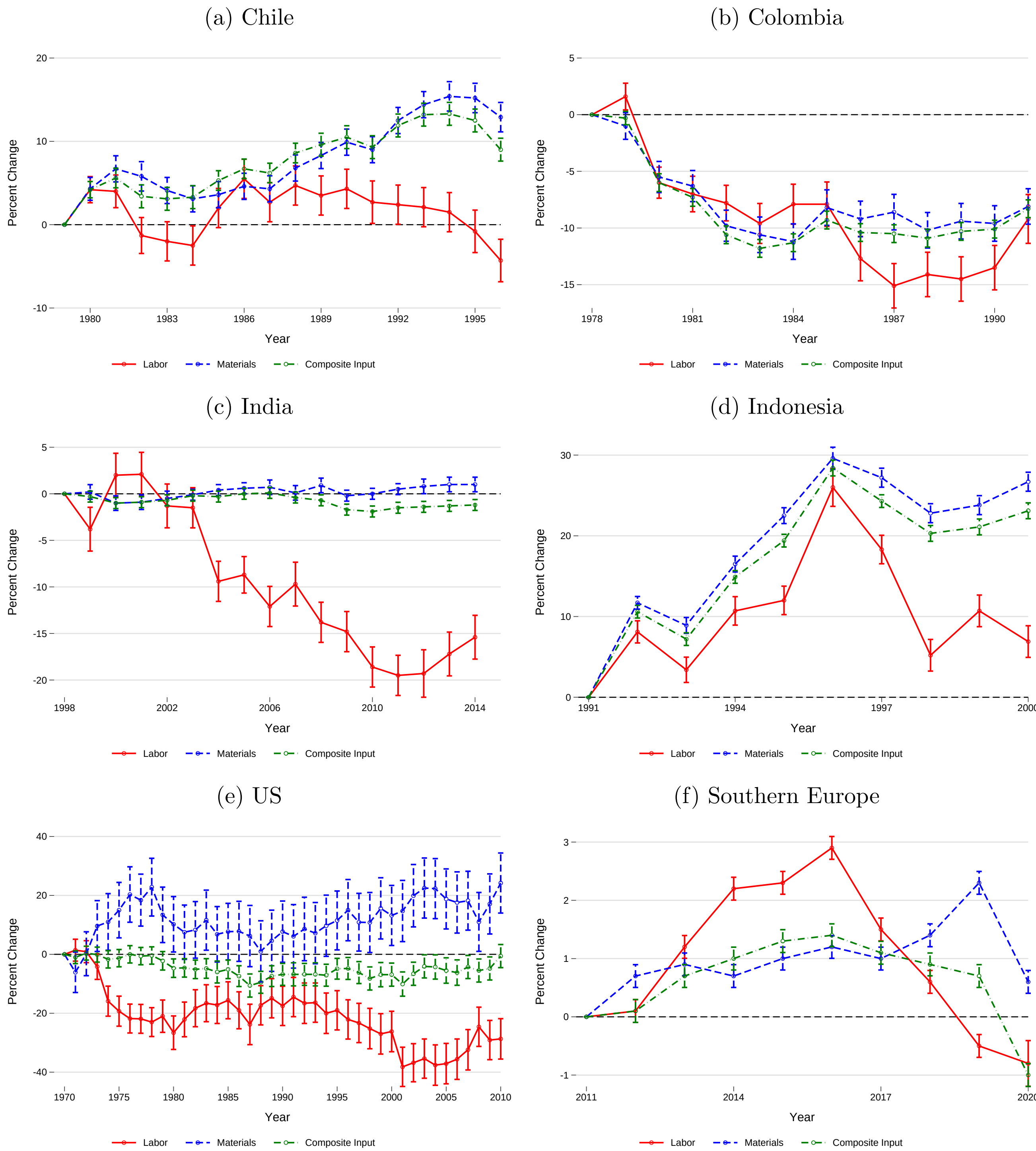
Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

Figure 7 Markup Time Trends using Flynn, Gandhi, Traina Estimates



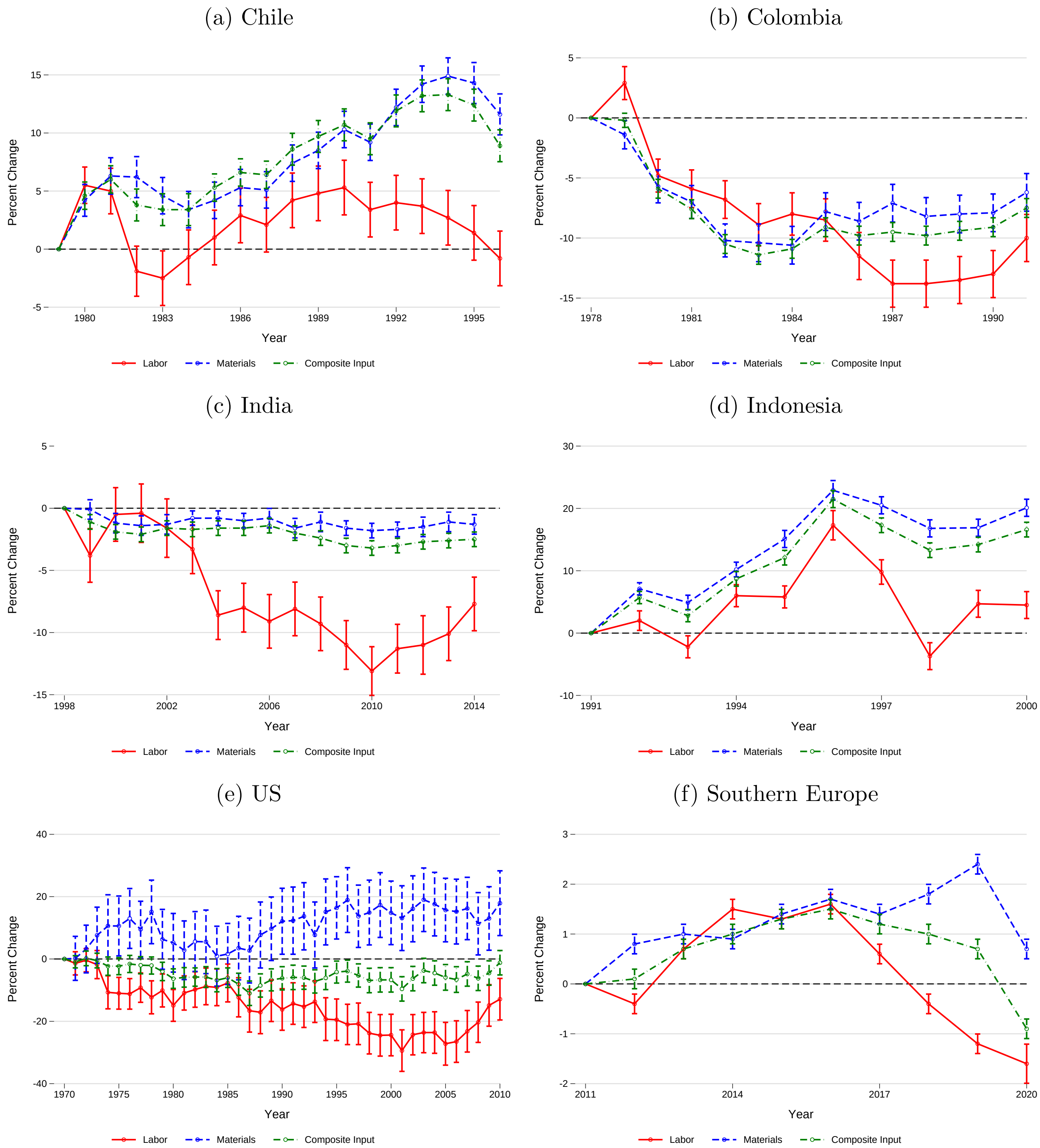
Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

Figure 8 Markup Time Trends using Industry Cost Share Estimates



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

Figure 9 Markup Time Trends using Subindustry Cost Share Estimates



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

I first examine this concern by estimating production functions at the subindustry level. There are 60 such subindustries for Chile, 82 for Colombia, 260 for Indonesia, 19 for the US (at NAICS 3 digit level, i.e. similar to the baseline industry definition for the other datasets), and 292 for Southern Europe. For India, industry definitions vary over time; there are 764 subindustries in the period before 2004, 684 between 2004 and 2007, and 586 in the period after 2007.[25]

I estimate production functions at the subindustry level using subindustry-year cost shares. Time trends, reported in Figure 9, continue to be very different across inputs. The magnitude of the negative cross-sectional correlation between the labor and materials markup is smaller at the subindustry level; the labor markup is uncorrelated with the materials markup for Indonesia, and is negatively correlated with the materials markup in the other datasets, with a 1% increase in the materials markup decreasing the labor markup by −0.20% to −1%. See the CostShare SubInd column of Table V.

For India, I also have access to product-level data and so can estimate product level production functions. I only include manufacturing plants that report only one product within a given year; in 2014, this dataset includes about 25,000 plants and 3,000 products. I then estimate production functions at the product-year level using product-year cost shares. The labor markup is negatively correlated with the materials markup, with a decline in the labor markup of −0.45% with a 1% increase in the materials markup using product-year cost shares, compared to −0.85% estimating production functions using industry-year cost shares on the same data.

Thus, estimating subindustry or product level production functions reduces, but does not eliminate, the negative cross-sectional correlation between markup estimates that I document.

C.3 Revenue Production Functions

Economists typically only have data on revenue, and not output, and so estimate revenue production functions. However, with imperfect competition, the markup is an additional unobservable in the revenue production. With imperfect competition, the control function estimator applied to revenue production functions may fail to identify production function parameters (Flynn et al., 2019; Doraszelski and Jaumandreu, 2019).

I examine this issue by using data on ten Indian homogenous products for which I have the quantity produced and price of the good, in the spirit of Foster et al. (2008).[26] For these products, I estimate product-level quantity production functions using the control function estimator. I only include plants for which at least 75% of their revenue comes from one of these products. The labor markup and materials markup are negatively correlated for these products, with a decline in the labor markup of −0.42% and −0.83% with a 1% increase in the materials markup using Cobb-Douglas and translog production functions. Thus, problems with revenue as opposed To

[25]For Chile and Colombia, the subindustry is defined at the four digit ISIC (Rev.2) level, for Indonesia at the five digit ISIC (Rev.2) level, and for India at the five digit NIC 98 level before 2004, five digit NIC 04 level between 2004 and 2007, and five digit NIC 08 level after 2007.

[26]I describe the construction of these products in Appendix G.7; they are Biri Cigarettes, Black Tea, Corrugated Sheet Boxes, Matches, Portland Cement, Processed Milk, Refined Sugar, Parboiled Non-Basmati Rice, Raw Non-Basmati Rice, and Shelled Cashew Nuts.

production functions alone cannot explain my findings.

C.4 Measurement Error

Another potential concern is measurement error in data on inputs due to survey collection. For example, manufacturing plants may not respond to all survey questions (White et al., 2016). However, the retailer's data is based on the internal records of the firm, and so should have very little measurement error compared to survey data. I find similar patterns using the retailer's data as I did in the manufacturing datasets.

Measurement error may be more of an issue for smaller, less sophisticated plants compared to large plants. All of my baseline estimates do not weight by size. I examine sales and cost weights, as in De Loecker et al. (2020) and Edmond et al. (2018) in Appendix E.4, and find qualitatively similar findings to the unweighted results.

Finally, for the Cobb-Douglas production function, the negative correlation between the labor markup and materials markup is driven by a negative correlation between the labor share of revenue and the materials share of revenue, as the output elasticities are industry-specific constants. For measurement error to account for this correlation, measurement errors in payroll would have to be negatively correlated with measurement errors in materials expenditure. It is unclear why this would be the case, except for the US Compustat data where they are functionally related.

D Markup Stylized Facts

I now examine several stylized facts, including how markups correlate with size, competition, exporting behavior, and an alternative profit share based markup. For each variable Z_{it}, I estimate the following regression specification:

$$\log(\mu_{it}^X) = \alpha + \beta Z_{it} + \gamma_t + \delta_n + \epsilon_{it} \tag{16}$$

where μ_{it}^X is the markup estimate for establishment i in year t using input X, and γ_t and δ_n are year and industry fixed effects.

Estimates of all of the stylized facts vary in sign and magnitude across inputs and datasets using both the Cobb-Douglas and translog control function estimators, and often conflict with predictions from theory.

D.1 Size

Multiple theories of variable markups (Atkeson and Burstein, 2008; Melitz and Ottaviano, 2008) predict markups increasing in firm size. I examine this prediction by estimating (16) regressing markups on the logarithm of deflated sales, and report these estimates in Table VI. With the Cobb-Douglas estimates, markups estimated using labor are substantially higher for bigger firms, while markups estimated using materials are negatively correlated with firm size in all datasets.

Using the translog estimates, this correlation is negative for materials for six of seven datasets and negative for labor for four datasets.

Table VI Markups and Sales

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	0.12	-0.03	-0.02	-0.00	0.01	0.00
	(0.005)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)
Colombia	0.16	-0.01	-0.07	-0.00	0.00	0.01
	(0.004)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
India	0.21	0.05	-0.02	-0.00	0.01	0.01
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Indonesia	0.20	0.04	-0.06	-0.03	0.01	0.01
	(0.003)	(0.003)	(0.001)	(0.001)	(0.000)	(0.000)
US	0.05	-0.01	-0.01	0.06	0.06	0.05
	(0.007)	(0.009)	(0.008)	(0.007)	(0.002)	(0.003)
S Europe	0.13	-0.03	-0.13	-0.05	-0.01	-0.01
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Retailer	0.31	0.09	-0.01	-0.02	0.03	-0.04
	(0.004)	(0.008)	(0.000)	(0.001)	(0.000)	(0.001)

Note: Estimates are based on (16) where the independent variable is deflated sales. Standard errors are clustered at the establishment level.

D.2 Exporting

Atkeson and Burstein (2008) and Melitz and Ottaviano (2008) also predict that exporters, being more productive than the typical firm, will have larger markups; De Loecker and Warzynski (2012) focused on this question. I examine this question using an indicator variable for whether the establishment exports using the manufacturing census datasets.[27] Table VII contains these estimates. The correlation of markups estimated using the Cobb-Douglas estimator with exporting are positive for Chile, Colombia, and Indonesia using labor and for Colombia and Indonesia using materials. Using the translog estimates, this correlation is negative for labor for two of the four datasets, and positive albeit with a small magnitude for materials in all of the datasets.

D.3 Profit Share Markups

An alternative method to estimate markups has been to use data on profits to measure the markup. Returns to scale (RTS) are equal to the markup multiplied by one minus the share of profits s_π, or

[27]For Chile, I only have exporter information for plants from 1990; for India, for plants from 2008.

Table VII Markups and Exporting

Dataset	Labor Cobb-Douglas	Labor Translog	Materials Cobb-Douglas	Materials Translog	Composite Input Cobb-Douglas	Composite Input Translog
Chile	0.07	-0.11	0.04	0.03	0.05	0.04
	(0.018)	(0.016)	(0.007)	(0.006)	(0.003)	(0.003)
Colombia	0.17	0.02	-0.04	0.03	0.04	0.04
	(0.016)	(0.014)	(0.009)	(0.004)	(0.003)	(0.003)
India	-0.03	-0.15	0.01	0.02	0.03	0.02
	(0.011)	(0.008)	(0.002)	(0.002)	(0.001)	(0.001)
Indonesia	0.28	0.05	-0.02	0.01	0.03	0.03
	(0.012)	(0.011)	(0.004)	(0.004)	(0.001)	(0.001)

Note: Estimates are based on (16) where the independent variable is an indicator for whether the establishment exports. Standard errors are clustered at the establishment level.

$RTS = \mu(1 - s_\pi)$. Thus, given constant returns to scale, one can invert the profit share to estimate the markup. We would expect this profit share based markup to be highly correlated with the production approach based markup.

I examine how production based markups correlate with the profit share based markup, estimating the profit share in two ways. First, as in Gutiérrez and Philippon (2016), I calculate the profit based markup as sales divided by total costs, where capital costs are measured through a user cost approach as the multiple of capital stocks and rental rates. Second, for the retailer, I have data on accounting profits measured as earnings before interest and taxes (EBIT) and so can calculate a profit based markup as sales divided by sales minus profits.

I then regress the log production based markup on the log profit share based markup using (16). I report these estimates in Table VIII. Using the Cobb-Douglas estimates, this correlation is negative for materials for three of eight datasets and negative for labor for four of eight datasets. Using the translog estimates, this correlation is negative for materials for three of eight datasets and negative for labor for seven of eight datasets.

D.4 Competition

One explanation for high markups is less competition. I examine how markups correlate with competition for the retailer using its own classification of the degree of competition. The retailer classifies each store as facing either Low, Medium, or High competition, and records the number of competitors for each store. I examine the competition band in this section in Table IX, and a discretized number of competitors in Table X.

With the Cobb-Douglas estimates, labor markups are slightly lower with mroe competition, and materials markups are slightly higher. With the translog estimates, labor markups are substantially (9%) lower on average with competition, while materials markups rise slightly.

Table VIII Production Markup Estimates and Profit Based Markup

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	-0.03	-0.06	0.37	0.35	0.09	0.08
	(0.016)	(0.014)	(0.010)	(0.009)	(0.003)	(0.003)
Colombia	-0.15	-0.16	0.01	0.05	-0.00	0.01
	(0.018)	(0.014)	(0.013)	(0.007)	(0.004)	(0.003)
India	0.21	-0.05	0.15	0.18	0.02	-0.01
	(0.010)	(0.008)	(0.003)	(0.004)	(0.001)	(0.001)
Indonesia	0.06	-0.09	-0.12	-0.09	-0.03	-0.04
	(0.011)	(0.011)	(0.006)	(0.005)	(0.002)	(0.002)
US	-0.01	-0.27	0.77	0.72	0.33	0.31
	(0.040)	(0.042)	(0.055)	(0.040)	(0.018)	(0.019)
S Europe	-0.41	-0.24	0.12	0.04	0.00	0.00
	(0.007)	(0.005)	(0.006)	(0.002)	(0.001)	(0.001)
Retailer	1.81	-0.09	-0.08	-0.01	0.15	-0.17
	(0.027)	(0.041)	(0.003)	(0.003)	(0.003)	(0.003)
Retailer (EBIT)	2.00	0.85	-0.09	-0.09	0.16	-0.16
	(0.028)	(0.045)	(0.003)	(0.004)	(0.003)	(0.003)

Note: Estimates are based on (16) where the independent variable is the profit share based markup. Standard errors are clustered at the establishment level. All profit based markups are through a factor cost based profit measure, except for the last row which is an accounting profit (EBIT) based measure.

Here, theory is not as clear cut. On the one hand, we might expect from canonical models of competition that markups would decline with competition. On the other hand, uniform or near-uniform pricing by many large retailers (DellaVigna and Gentzkow, 2017) might lead to the same markups across stores facing different competition, and the retailer's own data shows that it uses only a small number of pricing zones.

Table IX Markups and Competition

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Medium	-0.004	-0.016	0.000	-0.001	0.001	-0.004
	(0.004)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)
High	-0.003	-0.088	0.004	0.002	0.006	-0.014
	(0.006)	(0.009)	(0.001)	(0.001)	(0.000)	(0.001)

Note: Estimates are based on (16) where the independent variable is the company-derived measure of competition; all estimates are relative to a retail store facing Low Competition. Standard errors are clustered at the establishment level.

Table X Markup and Number of Competitors

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
2	0.006	0.024	-0.001	-0.003	0.000	-0.001
	(0.007)	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)
3	-0.002	0.013	-0.000	-0.002	0.000	-0.001
	(0.007)	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)
4	-0.002	0.007	-0.001	-0.003	0.001	-0.004
	(0.007)	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)
5-9	-0.005	-0.031	0.001	-0.002	0.003	-0.007
	(0.006)	(0.008)	(0.001)	(0.001)	(0.000)	(0.001)
10+	-0.003	-0.085	0.004	0.001	0.007	-0.015
	(0.009)	(0.013)	(0.001)	(0.001)	(0.001)	(0.001)

Note: Estimates are based on (16) and are relative to a retail store with 0-1 competitors. Standard errors are clustered at the establishment level.

E Additional Empirical Results

E.1 Trends over Time

In Figure 10, I depict aggregate markup trends based on labor, materials, or the combined input of both as flexible inputs estimated using Cobb-Douglas production functions. In Figure 11, I depict aggregate markup trends based on labor, raw materials, or energy as flexible inputs estimated using Cobb-Douglas production functions.

E.2 Markup Dispersion

In Table XI and Table XII, I report the 75/25 ratio and 90/10 ratio of markup estimates.

Table XI 75/25 Ratio of Markup Estimates

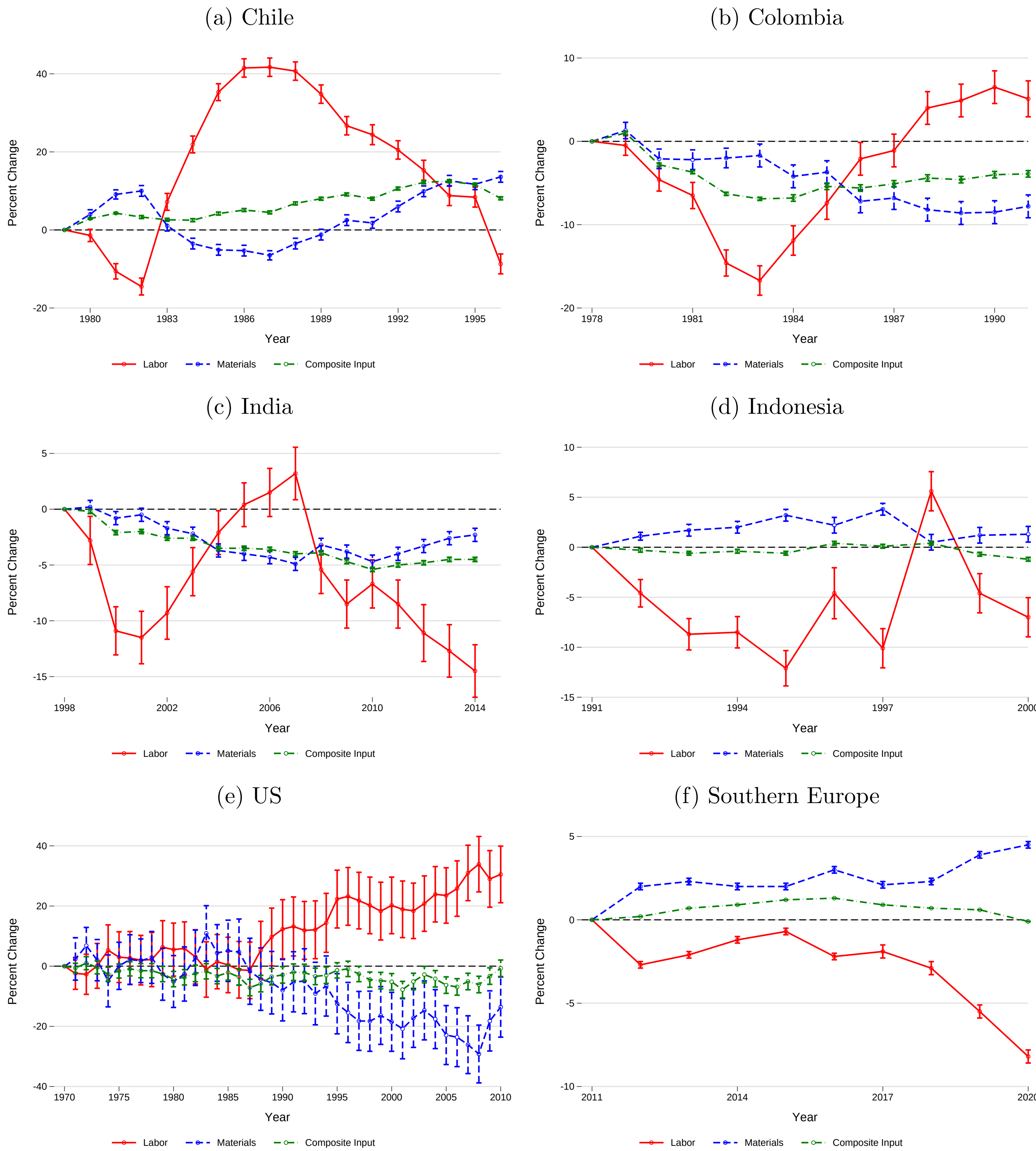
	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	2.68	2.06	1.41	1.32	1.16	1.15
	(0.009)	(0.010)	(0.003)	(0.003)	(0.001)	(0.001)
Colombia	2.69	1.87	1.63	1.24	1.14	1.14
	(0.013)	(0.006)	(0.005)	(0.001)	(0.001)	(0.001)
India	4.25	3.16	1.32	1.25	1.13	1.12
	(0.011)	(0.005)	(0.001)	(0.000)	(0.000)	(0.000)
Indonesia	3.82	2.65	1.55	1.37	1.12	1.13
	(0.022)	(0.010)	(0.002)	(0.002)	(0.000)	(0.000)
US	2.45	4.04	3.81	1.93	1.29	1.54
	(0.018)	(0.051)	(0.436)	(0.014)	(0.003)	(0.008)
S Europe	2.37	1.74	1.65	1.23	1.10	1.10
	(0.003)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Retailer	1.28	1.35	1.03	1.03	1.02	1.03
	(0.002)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Estimates use all establishments and years. Standard errors are based on 20 bootstrap simulations. For India, these estimates ignore the sample weights.

E.3 Average Markups

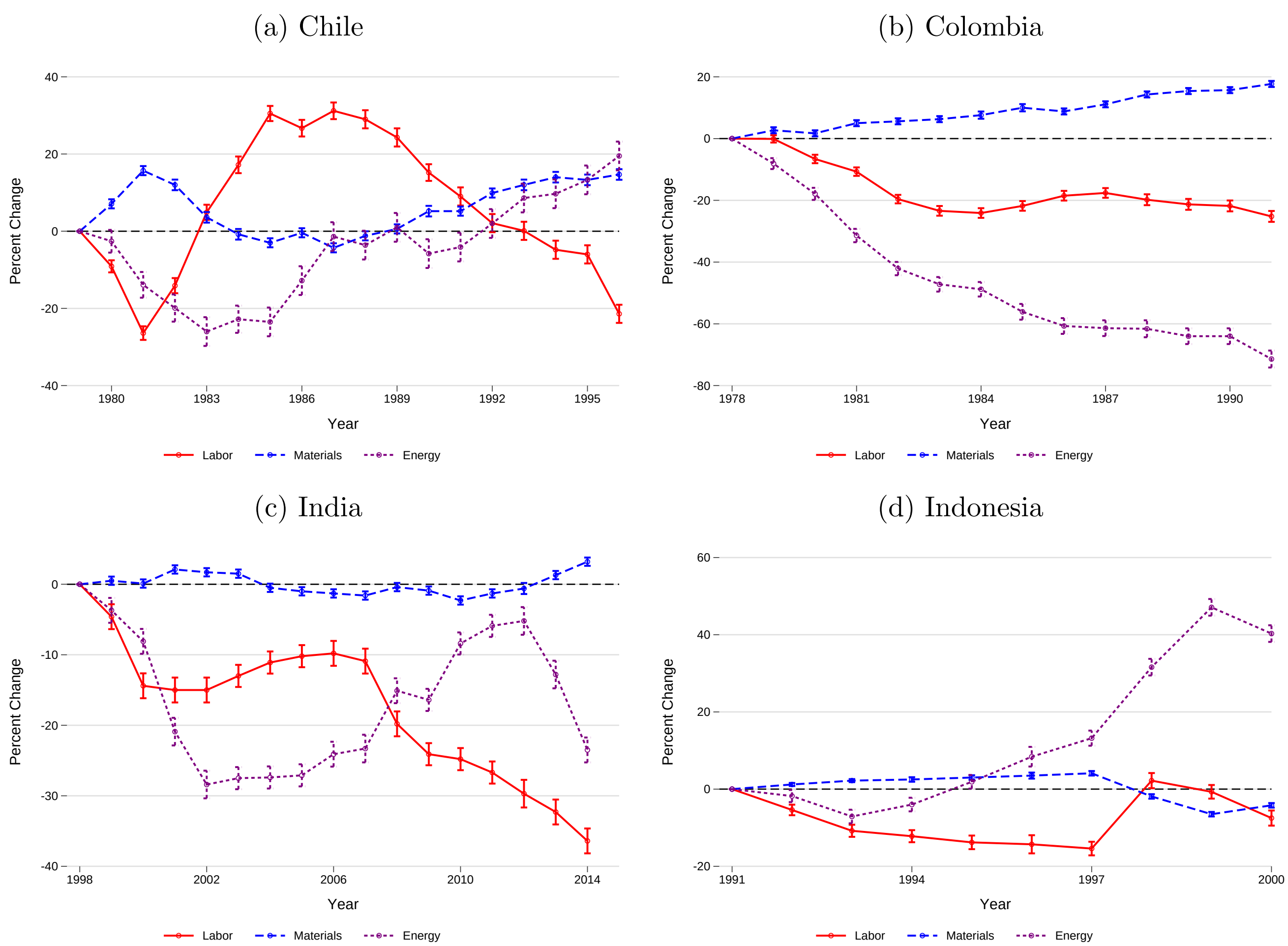
Under the production approach, the average markup should be the same using different flexible inputs. I test this prediction by estimating the average markup across all establishments using different flexible inputs. I find similar average markups in some, but not all, of the datasets.

Figure 10 Markup Time Trends using Cobb-Douglas Estimates



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

Figure 11 Markup Time Trends using Cobb-Douglas Estimates, with Energy



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero.

Table XII 90/10 Ratio of Markup Estimates

	Labor		Materials		Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	6.25	4.04	2.08	1.81	1.33	1.31
	(0.032)	(0.020)	(0.004)	(0.006)	(0.002)	(0.001)
Colombia	7.87	7.43	2.71	1.68	1.31	1.30
	(0.076)	(0.304)	(0.010)	(0.006)	(0.001)	(0.001)
India	15.81	10.08	1.75	1.58	1.27	1.27
	(0.063)	(0.044)	(0.001)	(0.001)	(0.000)	(0.000)
Indonesia	17.05	8.16	2.34	1.97	1.25	1.28
	(0.142)	(0.061)	(0.005)	(0.004)	(0.001)	(0.001)
US	6.42	14.51	-15.82	4.84	1.62	2.66
	(0.093)	(0.201)	(0.354)	(0.067)	(0.006)	(0.013)
S Europe	5.45	3.13	2.75	1.49	1.22	1.21
	(0.011)	(0.004)	(0.002)	(0.001)	(0.000)	(0.000)
Retailer	1.59	1.76	1.05	1.06	1.04	1.05
	(0.004)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Estimates use all establishments and years. Standard errors are based on 20 bootstrap simulations. For India, these estimates ignore the sample weights.

Using all the datasets, I report the ratio of the average labor markup to the average materials markup in the first two columns of Table XIII. The average labor markup is 9% higher than the average materials markup for Chile, 18% higher for Colombia, 98% higher for India, 72% higher for Indonesia, 137% higher for the US, 19% higher for Southern Europe, and 106% higher for the retailer under the Cobb-Douglas estimates. Under the translog estimates, the average labor markup is 50% higher than the average materials markup for Chile, 5% lower for Colombia, 46% higher for India, 69% higher for Indonesia, 124% higher for the US, 1% lower for Southern Europe, and 5% lower for the retailer. Thus, the average markups are close to each other for Colombia, Southern Europe, and the retailer using the translog estimates, and for Chile, Colombia, and Southern Europe using the Cobb-Douglas estimates.

E.4 Weighted Estimates

De Loecker et al. (2020) weight markups by sales, while Edmond et al. (2018) argue that cost weights are the right benchmark for welfare calculations. In this section, I weight all observations using sales weights (the plant's share of total sales in the year), or cost weights (the plant's share of total costs in the year). I then report the ratio of average markups, trends over time, and correlations between markups, using either labor, materials, or the combined variable input to compute markups. In some of the manufacturing datasets, a few plants have very large sales and cost shares (for example, petroleum refineries in India), so weighted estimates can differ from

Table XIII Ratio of Average Markup Estimates

	Labor/Materials		Labor/Composite Input		Materials/Composite Input	
Dataset	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog
Chile	1.09	1.50	1.30	1.63	1.19	1.09
	(0.012)	(0.012)	(0.012)	(0.012)	(0.003)	(0.002)
Colombia	1.18	0.95	1.53	1.02	1.30	1.08
	(0.016)	(0.015)	(0.016)	(0.013)	(0.010)	(0.005)
India	1.98	1.46	2.17	1.56	1.10	1.07
	(0.008)	(0.005)	(0.008)	(0.005)	(0.001)	(0.001)
Indonesia	1.72	1.69	2.00	1.89	1.17	1.11
	(0.018)	(0.019)	(0.019)	(0.021)	(0.003)	(0.002)
US	2.37	2.24	1.40	2.05	0.59	0.92
	(0.125)	(0.105)	(0.035)	(0.090)	(0.023)	(0.034)
S Europe	1.19	0.99	1.56	1.08	1.31	1.09
	(0.006)	(0.002)	(0.004)	(0.002)	(0.005)	(0.002)
Retailer	2.06	0.95	1.32	0.95	0.64	1.00
	(0.004)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)

Note: Estimates are the ratio of the average markup between two flexible inputs across all establishments and years, so Labor/Materials indicates the ratio of the average labor markup to average materials markup. Standard errors are clustered at the establishment level.

unweighted estimates substantially. Nevertheless, I continue to find negative correlations between labor markups and materials markups and different trends over time after weighting using sales or cost weights.

Table XIV Relationship between Markup Estimates: Sales Weighted

Dataset	Cobb-Douglas	Translog
Chile	-0.83	-0.30
	(0.060)	(0.076)
Colombia	-1.37	-0.09
	(0.087)	(0.199)
India	-1.89	-0.73
	(0.127)	(0.117)
Indonesia	-0.65	-0.30
	(0.094)	(0.111)
US	-0.21	-0.61
	(0.070)	(0.060)
S Europe	-1.47	-0.20
	(0.036)	(0.116)
Retailer	-7.06	-9.70
	(0.152)	(0.121)

Note: Estimates based on (6) where the labor markup is the dependent variable and materials markup the independent variable. Standard errors are clustered at the establishment level. Estimates weighted with sales weights.

F Monte Carlo

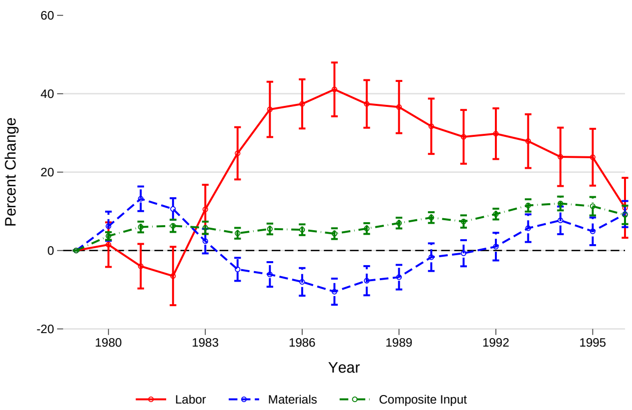
Through a Monte Carlo exercise, I show that labor augmenting productivity differences can cause a negative correlation between markups estimated using labor and materials as flexible inputs.

I simulate an economy in which markups and labor augmenting productivity differences vary across plants. In this economy, 1000 cost minimizing plants produce for 10 years. All plants have a common CES production function, as in (8), with substitution elasticity 0.5. The logarithm of neutral productivity A and labor augmenting productivity B evolve over time through an autoregressive process with a productivity persistence parameter of 0.9 and jointly normal shocks. Productivity is thus distributed as a joint lognormal. I then calibrate the parameters of this lognormal to match moments from data on factor shares and productivity from US manufacturing plants.[28]

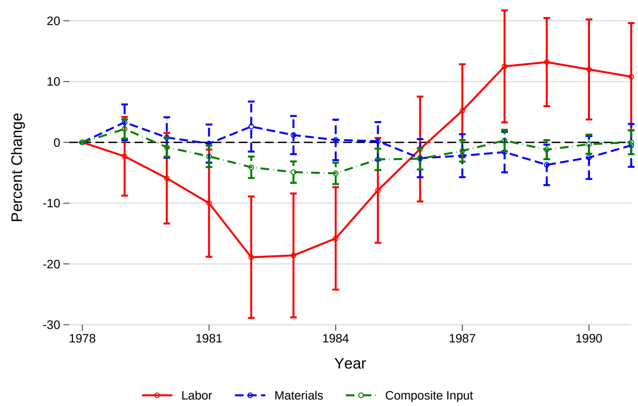
[28] I initialize productivities in their first year to the stationary distribution given the persistence process. I normalize the mean of the stationary distribution of $\log A$ to 1, and calibrate the mean of the stationary

Figure 12 Markup Time Trends using Cobb-Douglas Estimates, Sales Weighted

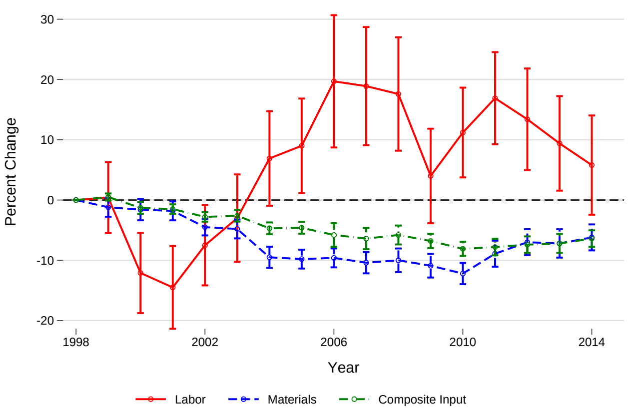
(a) Chile



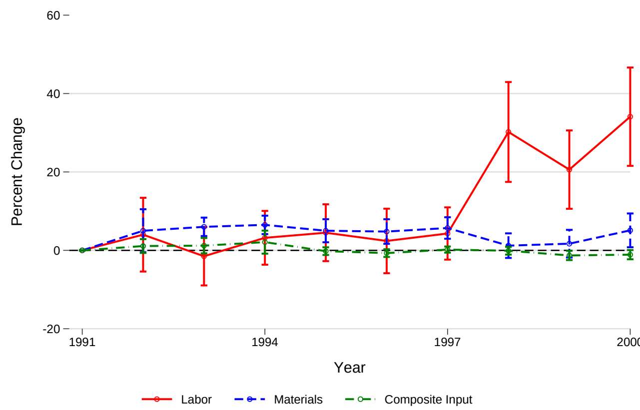
(b) Colombia



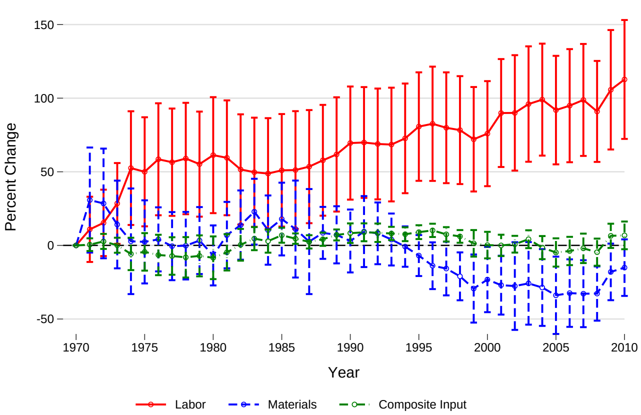
(c) India



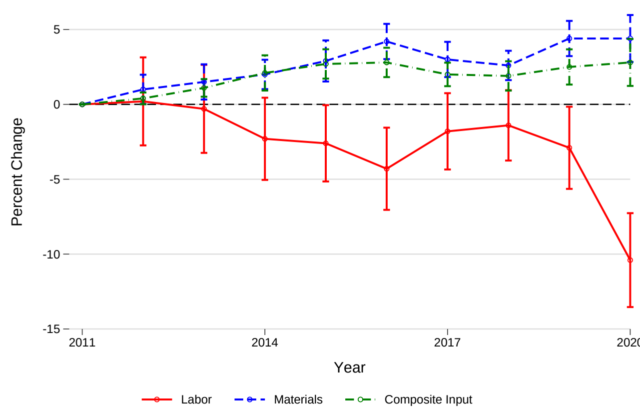
(d) Indonesia



(e) US



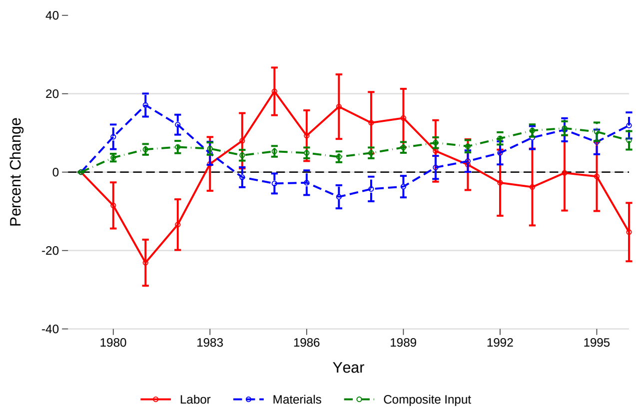
(f) Southern Europe



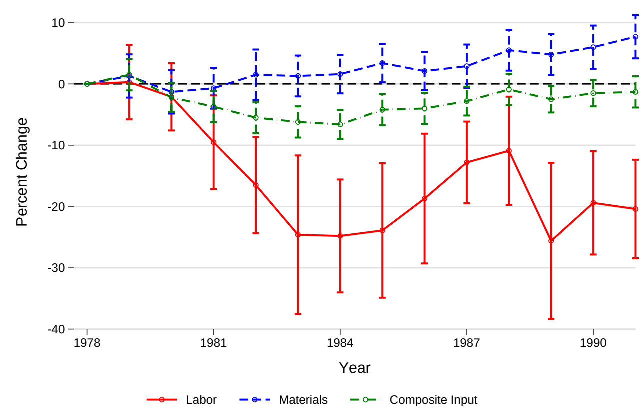
Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero. Estimates weighted with sales weights.

Figure 13 Markup Time Trends using Translog Estimates, Sales Weighted

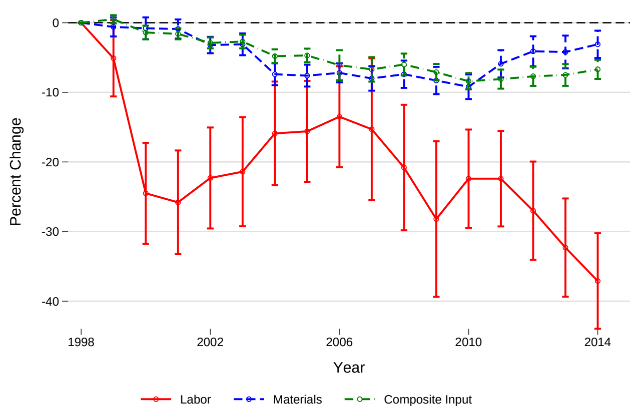
(a) Chile



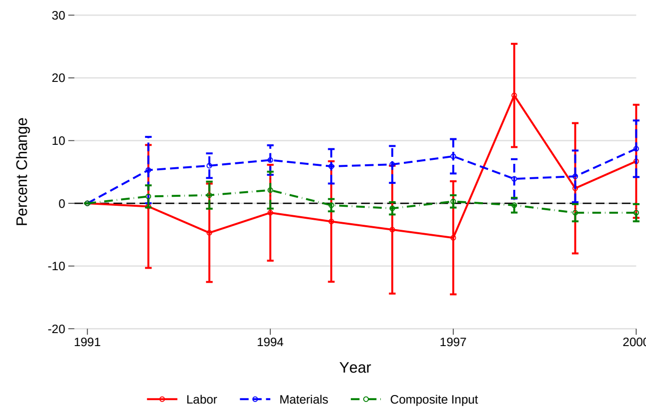
(b) Colombia



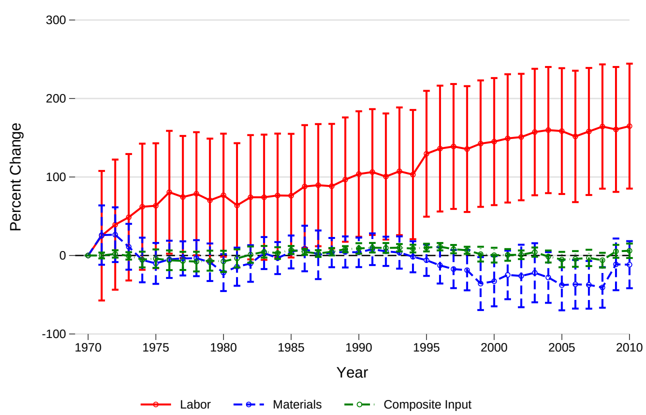
(c) India



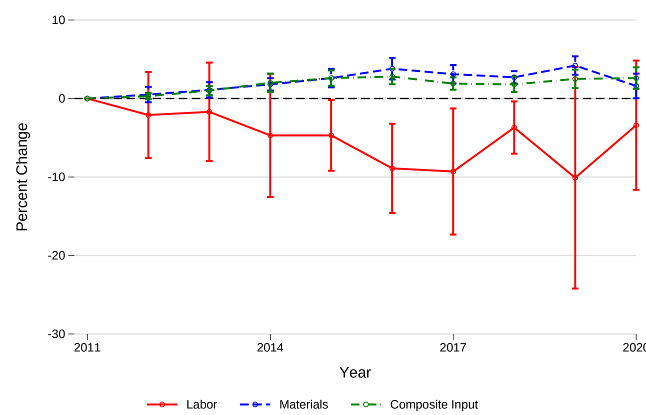
(d) Indonesia



(e) US

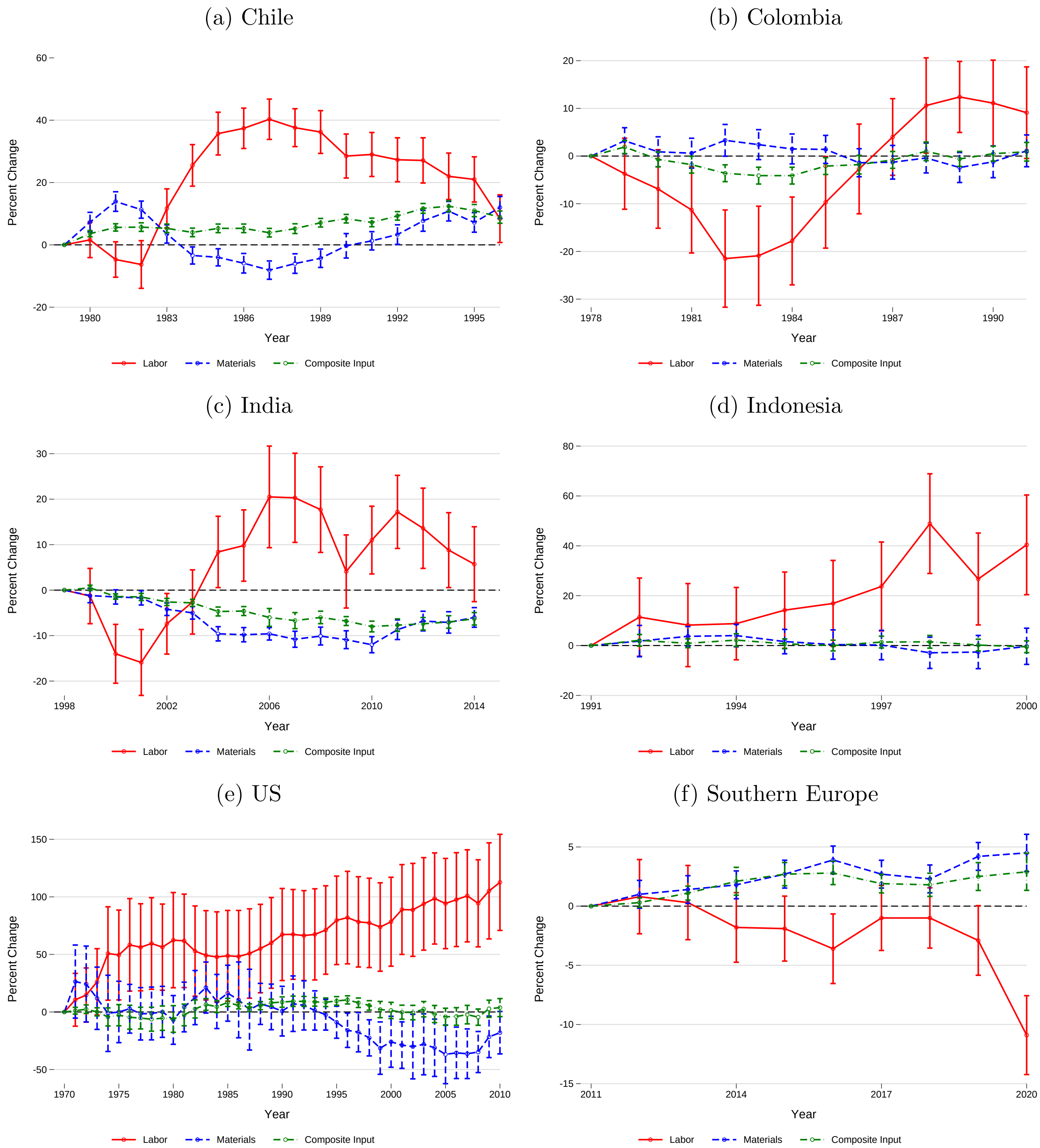


(f) Southern Europe



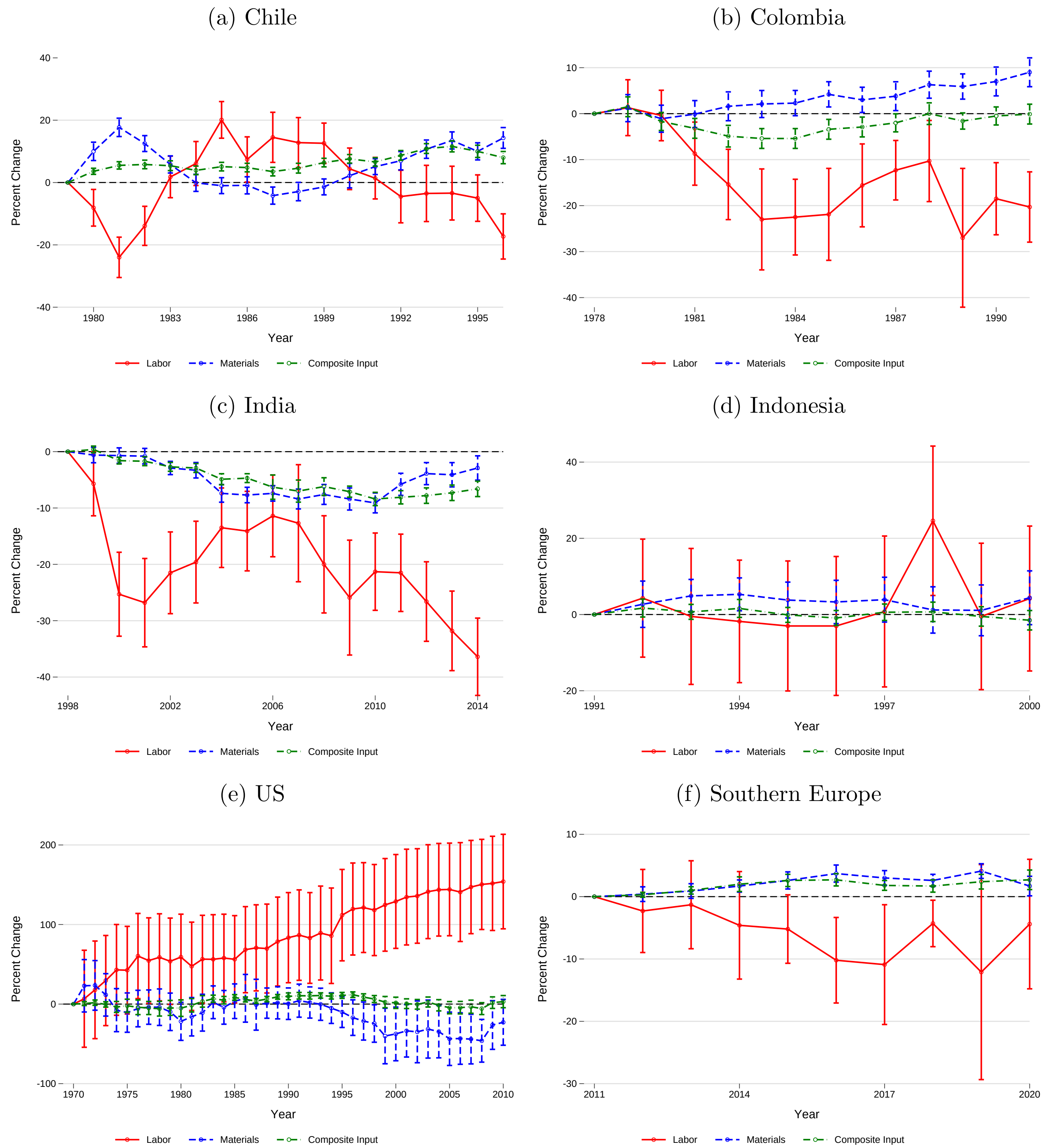
Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero. Estimates weighted with sales weights.

Figure 14 Markup Time Trends using Cobb-Douglas Estimates, Cost Weighted



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero. Estimates weighted with cost weights.

Figure 15 Markup Time Trends using Translog Estimates, Cost Weighted



Note: Estimates based on (5), and include 95% Confidence Intervals (vertical bars) based on clustering at the establishment level. All estimates relative to the first year, which is set to zero. Estimates weighted with cost weights.

Table XV Relationship between Markup Estimates: Cost Weighted

Dataset	Cobb-Douglas	Translog
Chile	-0.83	-0.29
	(0.059)	(0.069)
Colombia	-1.42	-0.08
	(0.068)	(0.161)
India	-1.98	-0.77
	(0.120)	(0.112)
Indonesia	-0.86	-0.46
	(0.116)	(0.126)
US	-0.23	-0.64
	(0.074)	(0.062)
S Europe	-1.53	-0.18
	(0.038)	(0.131)
Retailer	-7.07	-9.71
	(0.155)	(0.119)

Note: Estimates based on (6) where the labor markup is the dependent variable and materials markup the independent variable. Standard errors are clustered at the establishment level. Estimates weighted with cost weights.

Plants face CES demand with an elasticity of demand drawn from a uniform distribution between 2 and 6. Because demand is CES, the markup plants choose is a simple inversion of the demand elasticity; markups range between 1.2 and 2. Plants then set all inputs flexibly given the factor prices they face and their productivity draws.

I estimate the relationship between markup estimates using the following regressions:

$$\log(\mu_{it}^{L}) = \alpha + \beta \log(\mu_{it}^{M}) + \epsilon_{it} \tag{17}$$

$$\log(\mu_{it}^{True}) = \alpha + \beta \log(\mu_{it}^{X}) + \epsilon_{it}. \tag{18}$$

First, I compare the labor markup to the materials markup using (17). Second, I examine how the true markup based on the demand elasticity the plant faces is correlated with different production based markups for input X using (18). Here, the (logged) true markup is the dependent variable

distribution of $\log B$ and the variances and covariance of $\log A$ and $\log B$ through moment-matching. I match the following six moments: an aggregate capital share of capital and labor cost of 0.3, a value of the weighted variance of capital shares of capital and labor of 0.1, and the aggregate materials share of total cost of 0.55 (all from Oberfield and Raval (2021)) the 90-10 ratio of marginal cost across plants to 2.7 (from Syverson (2004)), the coefficient of a regression of the capital cost to labor cost ratio on the log of the plant's total cost of capital and labor (weighting by the plant's total cost of capital and labor) of 0.08 from Raval (2019), and a log of total industry cost of $\log(10,000)$ (to keep the same size industry across simulations). Distribution parameters are 0.1 for capital, 0.3 for labor, and 0.6 for materials.

and the labor, materials, or composite markup the independent variable.

In Table XVI, I report the average of β across 200 Monte Carlo simulations, with standard deviations across simulations in parentheses. I examine Cobb Douglas and translog control function estimators, for which B is assumed not to vary across plants.[29]

As I found in Section 3, labor markups are negatively correlated with materials markups. A 1% increase in the materials markup decreases the labor markup on average by 1.15% using the Cobb-Douglas control function estimator and 0.08% using the translog control function estimator.

In addition, both labor and materials markups are only slightly correlated with the true markup using the control function estimators; on average, the true markup is only 0.12% higher using the Cobb-Douglas estimator, or 0.05% higher using the translog estimator, after a 1% increase in the labor markup. The true markup is 0.36% higher using the Cobb-Douglas estimator, or 0.03% lower using the translog estimator, after a 1% increase in the materials markup.

Table XVI Relationship between Markup Estimates: Monte Carlo Estimates

Estimator	Labor on Materials	True Markup on Labor	True Markup on Materials	True Markup on Composite Input
Cobb-Douglas CF	-1.15	0.12	0.36	0.79
	(0.97)	(0.16)	(0.33)	(0.23)
Translog CF	-0.08	0.05	-0.03	0.27
	(0.33)	(0.10)	(0.02)	(0.21)

Note: Estimates based on 200 Monte Carlo simulations, using (17) and (18). For example, True Markup on Materials indicates a regression where the true markup is the dependent variable and materials markup the independent variable. True markup is the actual markup set by the firm based on its demand elasticity in the Monte Carlo simulations. Markup estimates are based on ACF control function estimators.

In all specifications, the composite input markup is more highly correlated with the true markup than labor or materials, as might be expected as the composite input combines two negatively correlated inputs. However, a 1% increase in the composite input markup increases the true markup by only 0.79% using the Cobb-Douglas estimates, and 0.27% using the translog estimates.

G Data Notes

In this section, I describe how I construct the main data variables for each dataset.

[29] The Cobb Douglas estimates are based on 114 of 200 simulations for labor and materials, and 197 of 200 simulations for the composite input, as in some simulations the coefficient on labor or materials was negative.

G.1 Datasets

The first dataset is the Chilean annual census of the manufacturing sector, Encuesta Nacional Industrial Anual (ENIA), spanning the years 1979 to 1996. This data covers all Chilean manufacturing plants with at least 10 employees, and so contains about 5,000 plants per year.

The second dataset is the annual Colombian Manufacturing census provided by the Departamento Administrativo Nacional de Estadistica between 1981 and 1991. This data contains about 7,000 plants per year. Plants with less than 10 employees are excluded in 1983 and 1984.

The third dataset is India's Annual Survey of Industries (ASI) from 1998 to 2014. Manufacturing establishments with over 100 workers are always sampled, while a rotating sample of one-third of all plants with at least ten workers (twenty if without power) are also sampled. I thus weight by the provided sample weights in samples using the Indian data. This data contains about 30,000 plants per year.

The fourth dataset is the Manufacturing Survey of Large and Medium-Sized Firms (Survei Industri, SI) from 1991 to 2000. This dataset is an annual census of all manufacturing firms in Indonesia with 20 or more employees, and contains about 14,000 firms per year.

The fifth dataset is Compustat, which comprises US public firms. I restrict the data to include only firms reporting a NAICS code for manufacturing between 1970 and 2010. This data contains about 500 firms per year, increasing from about 200 per year at the beginning of the sample to 1,000 per year at the end of the sample.

The sixth dataset is the ORBIS data from Bureau Van Dijk for manufacturing firms, defined as NACE Rev. 2 industry codes between 10 and 33, located in Italy, Spain, or Portugal ("Southern Europe"). The ORBIS data includes both public and private firms. Because my access to this data only includes firms active when I accessed the data, and up to 10 years of records for each firm, I only include the balanced panel of firms present in all years from 2011 to 2020. After data cleaning, this data contains about 100,000 firms per year, with about 30,000 firms per year in Spain, 54,000 firms per year in Italy, and 13,000 firms per year in Portugal.

I then compare total turnover and employment in this balanced panel to total turnover and employment for the manufacturing sector available from Eurostat. In 2019, the balanced panel comprises 67% of turnover and 56% of employment for Spain, 69% of turnover and 62% of employment for Italy, and 78% of turnover and 65% of employment for Portugal.[30]

The seventh dataset is store-level data from a major US nation-wide retailer for three years, comprising thousands of stores across the United States.

[30]Unlike the balanced panel, the unbalanced panel comprises a much lower share of turnover and employment at the beginning of the sample compared to the end of the sample. In 2011, the balanced panel comprises 65% of turnover and 51% of employment for Spain, 61% of turnover and 54% of employment for Italy, and 75% of turnover and 58% of employment for Portugal. The unbalanced panel is 93% of turnover and 89% of employment for Spain, 95% of turnover and 94% of employment for Italy, and 101% of turnover and 89% of employment for Portugal in 2019, and 78% of turnover and 69% of employment for Spain, 70% of turnover and 62% of employment for Italy, and 91% of turnover and 75% of employment for Portugal in 2011.

G.2 Capital

Capital costs are the most involved variable to construct. For each dataset, a capital stock is constructed for each type of capital available. Capital services is the sum of the stock of each type multiplied by its rental rate plus rental payments. This provides an approximation to a Divisia index for capital given different types of capital. See Diewert and Lawrence (2000) and Harper et al. (1989) for details on capital rental rates and aggregation.

For all datasets except the US, the capital rental rate is the sum of the real interest rate R and depreciation rate δ for that type of capital. I base the real interest rate on private sector lending rates reported in the World Bank World Development Indicators, which come from the IMF Financial Statistics, for each country. This real interest rate is constructed as the private sector lending rate adjusted for inflation using the change in the GDP deflator. Thus, real interest rate R is defined as $R = \frac{i_t - \pi_t}{1+\pi_t}$ for lending rate i_t and inflation rate π_t.

I average this real interest rate over the sample period, so that, since capital rental rates are constant over time, no variation in the capital stock over time is due to changing rental rates.[31]

For depreciation rates, I match the depreciation rates calculated for US industries to the equivalent industries in each country for structures and equipment. For transportation, I set the depreciation rate to 0.19.[32]

For the US, I use rental rates developed in Oberfield and Raval (2021) at the 3 digit NAICS level based on an external real rate of 3.5%, which account for the tax treatment of investment. Because these rental rates are separate for structures and equipment, I aggregate them using an average of the current and lag share of equipment in total capital for the industry from the NBER Productivity Database.

Across datasets, there are some differences in the construction of capital stocks. For Chile, I use end of year capital stocks constructed by Greenstreet (2007). Greenstreet (2007) constructed capital stocks for three types of capital – structures, equipment, and transportation – using a permanent inventory type procedure using data on capital depreciation.

For Colombia, India, and Indonesia, I construct asset-specific capital stocks using a perpetual inventory method for each type of capital. For Colombia, there are four types of capital: land, structures, equipment (combining office equipment and machinery), and transportation. For India, there are six types of capital: land, structures, equipment, transportation, computers, and other (including pollution equipment). For Indonesia, there are five types of capital: land, structures,

[31] For Chile and Colombia, the real interest rate series starts in 1985 and 1986, respectively, so I use interest rates starting from these dates. For Southern Europe, the World Development Indicators only include the real interest rate for Italy, so I use the averaged Italian rate for Spain and Portugal as well.

[32] The US depreciation rates are based on NIPA data on depreciation rates of assets; I then use asset-industry capital tables to construct depreciation rates for structures and equipment for each industry. Industries for the US are at the 2 digit SIC level. The US light truck depreciation rate is 19%. For the US and Southern Europe, I construct an average depreciation rate at the 3 digit NAICS level aggregating depreciation rates of structures and equipment using an average of the current and lag share of equipment in total capital from the NBER Productivity Database. I match by year for the US and use the 2010 depreciation rates for Southern Europe.

equipment, other capital (for which I use the equipment deflator), and transportation.[33]

For each asset type, I construct a perpetual inventory measure of capital starting with the first year reporting a positive value of the book value of capital. I also construct a backwards perpetual inventory measure of capital to create capital stocks for plants missing capital stocks using the forward perpetual inventory calculation.[34] I drop observations with zero or negative capital services for equipment or for total capital.

For the US, I have data on the aggregate net book value of plant, property, and equipment and yearly capital expenditures, but neither capital nor investment measures are broken out by type of asset. I first convert historical cost tangible fixed assets to current cost fixed assets using BEA historical cost to current cost deflators at the 3 digit NAICS level. I then construct a perpetual inventory measure of capital starting with the first year reporting a positive value of the book value of capital.

For Southern Europe, only the book value of tangible fixed assets is available, so I use the book value as the measure of capital.

Capital deflators for Chile and Colombia are at the 3 digit ISIC level, and I have separate deflators for structures, equipment, and transportation. For India, Indonesia, the US, and Southern Europe, I use a general capital deflator: at the 4 digit ISIC level for Indonesia, at the aggregate level for India, at the 3 digit NAICS level for the US (aggregated via a Tornqvist index with investment weights from 4 digit NAICS deflators available from the NBER productivity database), and at the 2 digit NACE level for Southern Europe.[35]

For the retailer, I have better data on capital than in the manufacturing datasets – the history of all investments by store going back to the early 1980s separately for land, structures, and equipment. I use this data to construct a perpetual inventory measure of capital for each type of capital. I obtain capital deflators and rental prices for each type of capital from the BLS Multifactor Productivity program, constructed for the retail trade industry.

Nominal capital services are then the sum of the real capital stock of each asset type multiplied by the appropriate deflator and capital rental rate, plus rent. Real capital services are the sum of the real capital stock of each asset type multiplied by the appropriate capital rental rate, plus deflated rent.[36]

[33] For other capital, I use the depreciation rate and deflator for equipment. For computers, I use a depreciation rate of 31.19%, the US depreciation rate for computer equipment.

[34] For Indonesia, only total capital and total investment are available in 1996. I thus restart the perpetual inventory capital measure in 1997, and the backwards PI measure in 1995.

[35] For Southern Europe, I obtain capital deflators from Eurostat. These deflators are at the country level with separate deflators for Spain, Portugal, and Italy. For Spain, these deflators are at the manufacturing level only, while for Italy, some NACE industries are combined. In particular, industries 16 through 18, 22 through 23, 24 through 25, 29 through 30, and 31 through 33 are combined for Italy.

[36] For Chile, rent is not differentiated by capital type, so I deflate using the structures deflator. Colombia differentiates between structures rent and machinery rent, India between land rent, building rent, and machinery rent (I use net rents for all three), and Indonesia between land rent and structures/machinery rent. For the US, rent is not differentiated by capital type, so I deflate using the overall investment deflator. For Southern Europe, data on rental payments are not available. For the retailer, I deflate rent using the structures deflator, as most capital is structures.

G.3 Labor

For Chile, Colombia, Indonesia, US, and Southern Europe, I use the total number of workers as my measure of labor. For India, I use the total number of days worked by all workers, while for the retailer, I use the total number of hours worked by all workers.

For labor costs, I use the sum of total salaries and benefits for all of the datasets except the US. For the US, data on wages are only available for about 10% of the sample. I thus follow Keller and Yeaple (2009) and Demirer (2020) and set labor costs as the number of employees multiplied by an average wage for the 3 digit NAICS industry. I measure the average wage from the NBER productivity database as payroll divided by employment aggregated to the 3 digit NAICS level, based upon a Tornqvist index with payroll based weights across 6 digit NAICS industries. Because the payroll data in the NBER Productivity database does not include benefits, I multiply this average wage by the ratio of compensation to wages and salaries at the 3 digit NAICS level available from the BEA.[37]

G.4 Energy and Materials

I can separate energy costs from raw materials for the four manufacturing censuses. Total energy costs are expenses on all energy inputs, subtracting out any electricity sold to other parties.

Real energy input requires energy deflators. For Chile, I have data on both value and quantity of energy inputs for 10 different inputs (plus other fuel). I follow Greenstreet (2007)'s construction of deflators for each energy input as the ratio of total value over total quantity for each 3 digit industry-year. Other fuel is deflated using a value weighted average of the other fuels. Electricity is deflated calculating an electricity price as the average total value of electricity over total quantity for the year.

For Colombia, I calculate the average electricity price as the median ratio of value to quantity across all plants for a given year and province and deflate electricity using this electricity price. For fuels, I only have aggregate fuel value, which I deflate using the output deflator for the 3 digit petroleum and coal industry.

For India, I deflate fuels and electricity using yearly deflators for each input.

For Indonesia, I calculate the average electricity price as the median ratio of value to quantity across all plants for a given year and deflate electricity using this electricity price. For fuels, I have data on both value and quantity of energy inputs for 7 different inputs (plus other fuel). I thus create deflators for each energy input based on the median value to amount ratio by year. I use the diesel oil deflator for other fuel inputs.

For Chile, Colombia, and India, I calculate total raw materials as total spending on raw materials, with an adjustment for inventories of raw materials by adding the difference between the end year and beginning year value of inventories of raw materials. For Indonesia, total amount of raw materials used are reported, which I use for total raw materials.

[37]For earlier years when only SIC data is available, I match SIC 2 digit industries to the equivalent NAICS 3 digit industries, and adjust for the industry difference using the average difference in ratio for 1998-2000 when both SIC and NAICS level data are available.

For Chile and Colombia, materials deflators are at the 3 digit ISIC level. For Indonesia, they are at the 5 digit ISIC level and for India at the 4 digit NIC 2008 level. For Chile, I also deflate lubricants, water, and grease using value to quantity ratios as for the energy inputs described above, following Greenstreet (2007). For Indonesia, I also do the same for lubricants.

Unlike all other datasets, for the US I do not have materials costs as a separate data field. Thus, I follow Keller and Yeaple (2009) and Demirer (2020) and define materials costs as cost of goods sold plus selling, general, and administrative expenses minus depreciation and minus labor costs as defined above. I deflate these materials costs using 3 digit NAICS level materials deflators aggregated via a Tornqvist index with materials cost weights from the NBER Productivity database.

For Southern Europe, I use materials costs for materials. I deflate materials by creating materials deflators at the NACE 2 digit level; these are a harmonic index of output deflators from Eurostat for the Euro 19 area at the NACE 2 digit level, with weights as shares from the 2015 product by product input-output table.

For the retailer, materials are the total cost of goods sold at the store. Real materials are constructed by deflating goods using the appropriate deflators from the PPI.

G.5 Sales

For the manufacturing censuses, I calculate total sales as total production value (both domestic sales and exports, and sales to other establishments of the same company), plus the difference between the end year and beginning year value of inventories of finished goods. Real sales are nominal sales deflated by the output deflator. The output deflator is measured at the 3 digit ISIC level in Chile and Colombia, at the 4 digit NIC 08 level in India, and the 5 digit ISIC level in Indonesia.

For the US, I use total sales and deflate using 3 digit NAICS level output deflators from the NBER Productivity database, aggregated via a Tornqvist index with sales weights. For Southern Europe, I use total sales and deflate using output deflators at the 4 digit NACE level (3 digit or 2 digit NACE if 4 digit deflator not available) for the Euro 19 area from Eurostat. For the retailer, I deflate total sales using PPI deflators for the relevant goods.

G.6 Industry Sectors and Data Cleaning

For Indonesia, I drop all duplicated observations. The industry definition also changes in 1998 from ISIC rev.2 to ISIC rev. 3 (with both reported in 1998). I assign plants in 1999 and 2000 the reported ISIC rev. 2 industry in 1998 if they exist in 1998; if not, I use the modal 5 digit ISIC rev.2 given the reported value of ISIC rev. 3 using data from 1998.

For India, the industry definition repeatedly changes over the sample period. I use the panel structure of the data to create a consistent industry definition at the NIC 08 level. For plants with a NIC 98 or NIC 04 industry, I set the plant's industry to either the modal industry at the NIC 08 level across years for the plant, or, if this fails, the modal industry at the NIC 08 level for the given NIC 04 or NIC 98 industry.

For both India and Indonesia, I follow Alcott et al. (2015) and drop plants with an electricity share of sales above one and a labor, materials, or energy share of sales above two, or sales below 3 currency units.

For the US, I only include firms incorporated in the US. Because firms vary in the end of their fiscal year, I set the year to be the current year for statements ending after May 31st and to the previous year otherwise. I remove firms with missing or non-positive values of sales, employment, cost of goods sold, selling, general, and administrative expenses, and depreciation. I also remove firms with missing capital expenditures and with less than 10 employees.

For Southern Europe, I only include firms incorporated in the relevant country (i.e. Italy for Italy, etc.). Because firms vary in the end of their fiscal year, I set the year to be the current year for statements ending after May 31st and to the previous year otherwise. I remove observations missing a BVD ID or year, as well as duplicate observations with the same BVD ID and year. I remove firms with missing or non-positive values of sales, employment, fixed assets, operating expenses, materials costs, and wages, or employment above 2 million employees. Finally, I only use the balanced panel, i.e. firms present for all 10 years of the sample.

G.7 Products

I construct ten homogeneous products in the Indian data. When doing so, I have to account for the fact that the product coding changes several times over the sample period. I describe each product below.

Biri cigarettes are recorded in thousands of cigarettes. In the 1998 to 2007 data, I use ASICC code 15323. In the 2008 to 2009 data, I use ASICC code 15325. In the 2010 to 2014 data, I use ASICC code 2509001.

Black Tea is recorded in kilograms. I include several product codes that correspond to black tea, but exclude non-black tea, tea bags, and instant tea. In the 1998 to 2009 data, I use the following ASICC codes: ASICC code 12211 [tea (black) leaf (blended)], ASICC code 12212 [tea (black) leaf (unblended)], ASICC code 12213 [tea (black) dust (blended)], ASICC code 12214 [tea (black) dust (unblended)], and ASICC code 12215 [tea (black) leaf (darjeeling)]. In the 2010 to 2014 data, I use the following ASICC codes: ASICC code 2391301 [Black Tea (CTC) "crush, tear, curl"], ASICC code 2391302 [darjeeling tea black leaf], ASICC code 2391303 [non-darjeeling black leaf], and ASICC code 2391308 [tea dust].

Boxes, Corrugated Sheet are recorded in number of boxes. In the 1998 to 2009 data, I use ASICC code 57104. In the 2010 to 2014 data, I use ASICC code 3215301.

Matches are recorded in kilograms. In the 1998 to 2009 data, I use ASICC code 37304. In the 2010 to 2014 data, I use ASICC codes 3899801 [Matches safety (match box)] and 3899899 [Matches n.e.c.].

Portland Cement is recorded in tonnes. In the 1998 to 2007 data, I use ASICC code 94415. In the 2008 to 2009 data, I use ASICC code 94414. In the 2010 to 2014 data, I use ASICC code 3744008.

Processed Milk is recorded in fluid liters. In the 1998 to 2009 data, I use the following ASICC codes: ASICC code 11401 [fresh milk], ASICC code 11402 [flavored milk], ASICC code 11403 [chilled/frozen milk], and ASICC code 11404 [skimmed/pasteurized milk]. In the 2010 to 2012 data, I use ASICC code 2211000 [processed liquid milk]. In the 2013 to 2014 data, I use the following ASICC codes: ASICC code 2211001 [full cream milk], ASICC code 2211002 [toned milk], ASICC code 2211003 [skimmed milk], and ASICC code 2211099 [other processed milk (nec)].

Refined Sugar is recorded in tonnes. In the 1998 to 2009 data, I use ASICC code 13103. After 2009, refined sugar is initially split into multiple codes with different units (kilograms vs. tonnes), so I do not include refined sugar after 2009.

Rice, Parboiled Non-Basmati is recorded in tonnes. In the 1998 to 2009 data, I use ASICC code 12311. In the 2010 to 2014 data, I use ASICC codes 2316107 [Rice (other than basmati), par-boiled milled] and 2316202 [Rice (other than basmati), par-boiled brown/ husked].

Rice, Raw Non-Basmati is recorded in tonnes. In the 1998 to 2009 data, I use ASICC code 12312. In the 2010 to 2014 data, I use ASICC codes 2316108 [Rice (other than basmati), non-boiled (atap) milled] and 2316203 [Rice (other than basmati), non-boiled (atap) brown/ husked].

Shelled Cashew Nuts is recorded in tonnes. In the 1998 to 2007 data, I use ASICC code 12111. In the 2008 to 2009 data, I use ASICC code 12131. In the 2010 to 2014 data, I use ASICC code 2142400.

I only keep manufacturing plants with a 75% of greater revenue share of a given product. I define the price of a product as the gross value of the product minus any reported expenses (excise duty, sales tax, and other expenses) divided by the quantity sold. I then drop all plants whose price is greater than five times, or less than 20%, of the median price for a given product in a given year.

Table XVII below contains the total number of observations, and number of distinct manufacturing plants, for each product.

Table XVII Homogeneous Products

Product	Number of Observations	Number of Distinct Plants
Biri Cigarettes	3234	1053
Black Tea	7263	1316
Boxes, Corrugated Sheet	4234	2299
Matches	2725	676
Portland Cement	2262	598
Processed Milk	2143	784
Refined Sugar	3612	600
Rice, Parboiled Non-Basmati	6433	4481
Rice, Raw Non-Basmati	5535	4061
Shelled Cashew Nuts	3118	979