Micro Data and Macro Technology*

Ezra Oberfield

Princeton University

ezraoberfield@gmail.com

Devesh Raval

Federal Trade Commission

devesh.raval@gmail.com

May 27, 2014

Abstract

We develop a framework to estimate the aggregate capital-labor elasticity of substitution by aggregating the actions of individual plants, and use it to assess the decline in labor's share of income in the US manufacturing sector. The aggregate elasticity reflects substitution within plants and reallocation across plants; the extent of heterogeneity in capital intensities determines their relative importance. We use micro data on the cross-section of plants to build up to the aggregate elasticity at a point in time. Our approach thus places no assumptions on the evolution of technology, so we can separately identify shifts in technology and changes in response to factor prices. We find that the aggregate elasticity for the US manufacturing sector has been stable since 1970 at about 0.7. Mechanisms that work solely through factor prices cannot account for the labor share's decline. Finally, we examine how the aggregate elasticity varies across countries and find it is substantially higher in less-developed countries.

KEYWORDS: Elasticity of Substitution, Aggregation, Labor Share, Bias of Technical Change.

*First version: January 2012. We thank Enghin Atalay, Bob Barsky, Ehsan Ebrahimy, Ali Hortacsu, Konstantin Golyaev, Sam Kortum, Lee Lockwood, Harry Paarsch, Richard Rogerson, Chad Syverson, and Hiroatsu Tanaka for their comments on this paper. We would also like to thank Frank Limehouse for all of his help at the Chicago Census Research Data Center. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or the Federal Trade Commission and its Commissioners. All results have been reviewed to ensure that no confidential information is disclosed.

1 Introduction

Over the last several decades, labor's share of income in the US manufacturing sector has fallen by more than 15 percentage points. A variety of mechanisms have been proposed to explain declining labor shares.[1] Some of these mechanisms, such as declining capital prices or increased capital accumulation, alter the labor share solely through changing factor prices. Others, such as automation and offshoring, would be viewed through the lens of an aggregate production function as a change in technology. As Hicks (1932) first pointed out, the crucial factor in assessing the relevance of these mechanisms is the aggregate capital-labor elasticity of substitution, which governs how aggregate factor shares respond to changing factor prices.[2] More generally, as a central feature of aggregate technology, the elasticity helps answer a wide variety of economic questions. These include, for example, the welfare impact of corporate tax changes, the impact of international interest rate differentials on output per worker, the speed of convergence to a steady state, and how trade barriers shape patterns of specialization.[3]

Unfortunately, obtaining the elasticity is difficult; Diamond et al. (1978) proved that the elasticity cannot be identified from time series data on output, inputs, and marginal products alone. Instead, identification requires factor price movements that are independent of the bias of technical change. Economists have thus explored two different approaches to estimating the elasticity.

The first approach uses the aggregate time series and places strong parametric assumptions on the aggregate production function and bias of technical change for identification. The most common assumptions are that there has been no bias or a constant bias over time. Leon-Ledesma

[1] See Karabarbounis and Neiman (2013), Piketty (2014), and Elsby et al. (forthcoming).

[2] Formally, changes in factor compensation can be decomposed into the response to shifts in factor prices (holding technology fixed) and changes due to shifts in technology (holding factor prices fixed). We call the latter the "bias of technical change," and we use the term loosely to include everything other than factor prices, including changes in production possibilities or in composition due to changing tastes or trade.

[3] See Hall and Jorgenson (1967), Mankiw (1995), and Dornbusch et al. (1980).

et al. (2010) demonstrated that, even under these assumptions, it is difficult to obtain the true aggregate elasticity. Not surprisingly, the estimates found in this literature vary widely.[4]

The second approach uses micro production data with more plausibly exogenous variation in factor prices, and yields the micro capital-labor elasticity of substitution. However, Houthakker (1955) famously showed that the micro and macro elasticities can be very different; an economy of Leontieff micro units can have a Cobb-Douglas aggregate production function.[5] Given Houthakker's result, it is unclear whether the micro elasticity can help answer the many questions that hinge on the aggregate elasticity.

In this paper, we show how the aggregate elasticity of substitution can be recovered from the plant-level elasticity. Building on Sato (1967), we show that the aggregate elasticity is a convex combination of the plant-level elasticity of substitution and the elasticity of demand.[6] In response to a wage increase, plants substitute towards capital and capital-intensive plants gain market share from labor-intensive plants. The degree of heterogeneity in capital intensities determines the relative importance of the within-plant substitution and reallocation.[7] For example, when all plants produce with the same capital intensity, there is no reallocation of resources across plants.

Using this framework, we build the aggregate capital-labor elasticity from its individual components. We estimate micro production and demand parameters. Since Levhari (1968), it is well known that Houthakker's result is sensitive to the distribution of capital intensities. Rather than making distributional assumptions, we directly measure the empirical distribution using cross-

[4]While Berndt (1976) found a unitary elasticity of substitution in the US time series assuming neutral technical change, Antras (2004) and Klump et al. (2007) subsequently found estimates from 0.5 to 0.9 allowing for biased technical change. Karabarbounis and Neiman (2013) estimate an aggregate elasticity of 1.25 using cross country panel variation in capital prices, while Piketty (2014) estimates an aggregate elasticity between 1.3 and 1.6; neither controls for biased technical change.

[5]Houthakker assumed that factor-augmenting productivities follow independent Pareto distributions. The connection between Pareto distributions and a Cobb-Douglas aggregate production function is also emphasized in Jones (2005), Lagos (2006), and Luttmer (2012).

[6]Sato (1967) showed this for a two-good economy. See also Miyagiwa and Papageorgiou (2007).

[7]We show that a statistic proportional to a cost-weighted variance of capital shares is sufficient to capture all of the micro heterogeneity.

sectional micro data.

Thus, given the set of plants that existed at a point in time, we estimate the aggregate elasticity of substitution *at that time*. Our strategy allows both the elasticity and the bias of technical change to vary freely over time, opening up a new set of questions. Because our identification does not impose strong parametric assumptions on the bias, our approach is well suited for measuring how it has varied over time and how it has contributed to the decline in labor's share of income. We can also examine how the aggregate elasticity has changed over time and whether it varies across countries at different stages of development.

We first estimate the aggregate elasticity for the US manufacturing sector using the US Census of Manufactures. In 1987, our baseline year, we find an average plant-level elasticity of substitution of roughly one-half. Given the heterogeneity in capital shares and our estimates of other parameters, the aggregate elasticity in 1987 was 0.7. Reallocation thus accounts for roughly one-third of substitution.[8] Despite large structural changes in manufacturing over the past forty years,[9] the aggregate elasticity has been stable. We find the elasticity rises slightly from 0.67 in 1972 to 0.73 in 2007.

We then use this estimate to decompose the decline in labor's share of income in the manufacturing sector since 1970. We have several findings. First, changes in factor prices have played only a small role in the decline in labor's share. Given our estimated aggregate elasticity, increased capital accumulation or declining capital prices would raise the labor share[10], while stagnating real wages can account for about one sixth of the decline. Second, the bias of technical change within industries has increased and accounts for most of the decline in the labor share. Third, a shift

[8]Our estimate is a long run elasticity of substitution between capital and labor, owing to a proper interpretation of our estimates of plant-level parameters. It is thus an upper bound on the short run elasticity.

[9]Elsby et al. (forthcoming) have emphasized the shift in industry composition in the context of labor's share of income.

[10]These mechanisms, put forward by Piketty (2014) and Karabarbounis and Neiman (2013) respectively, require an aggregate elasticity above one.

in the composition of industries accounts for part of the acceleration in the labor shares decline since 2000. These findings suggest the decline in the labor share stems from factors that affect technology, broadly defined, including automation and offshoring, rather than mechanisms that work solely through factor prices.

Finally, our approach allows us to examine how the shape of technology varies across countries at various stages of development; policies or frictions that lead to more variation in capital shares raise the aggregate elasticity. Using their respective manufacturing censuses, we find an average manufacturing sector elasticity of 0.81 for Chile, 0.84 for Colombia, and 1.15 for India. These differences are quantitatively important; the response of output per worker to a change in the interest rate is over fifty percent larger in India than in the US, as is the welfare cost of capital taxation. They also imply that a decline in capital prices decreases the labor share in India but increases it in the US.

Our work complements the broad literature that examines the shape of aggregate technology and the distribution of income. Krusell et al. (2000) study how a capital-skill complementarity and a declining price of capital can change factor shares and raise the skill premium. Acemoglu (2002), Acemoglu (2003), and Acemoglu (2010) show how factor prices and the aggregate capital-labor elasticity of substitution can determine the direction of technical change. Burstein et al. (2014) study how changes in technology and supplies of various factors impact factor compensation. Autor et al. (2003) and Autor et al. (forthcoming) study how changes in technology and trade impact factor income. Piketty (2014) examines the accumulation and distribution of capital.

The rest of the paper is organized as follows. In Section 2, we present our theoretical analysis of the aggregation problem, while in Section 3 we estimate the aggregate elasticity for manufacturing. Section 4 examines the robustness of our estimates. In Section 5, we use our estimates of the aggregate elasticity to assess the changes in the labor share. In Section 6, we examine how the aggregate elasticity varies across countries and discuss policy implications. Finally, in Section 7 we conclude.

2 Theory

This section characterizes the aggregate elasticity of substitution between capital and labor in terms of production and demand elasticities of individual plants. We begin with a simplified environment in which we describe the basic mechanism and intuition. We proceed to enrich the model with sufficient detail to take the model to the data by incorporating materials and allowing for heterogeneity across industries.

A number of features—adjustment costs, an extensive margin, non-constant returns to scale, and imperfect pass-through—are omitted from our benchmark model. We postpone a discussion of these until Section 4.

2.1 A Simple Example

Consider a large set of plants, I, whose production functions share a common, constant elasticity of substitution between capital and labor, σ. A plant produces output Y_i from capital K_i and labor L_i using the following CES production function:

$$Y_i = \left[(A_i K_i)^{\frac{\sigma-1}{\sigma}} + (B_i L_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{1}$$

Productivity differences among plants are factor augmenting: A_i is i's capital-augmenting productivity and B_i its labor-augmenting productivity.

Consumers have Dixit-Stiglitz preference across goods, consuming the bundle $Y = \left(\sum_{i\in I} D_i^{\frac{1}{\varepsilon}} Y_i^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$.

Plants are monopolistically competitive, so each plant faces an isoelastic demand curve with a common elasticity of demand $\varepsilon > 1$.

Among these plants, aggregate demand for capital and labor are defined as $K \equiv \sum_{i \in I} K_i$ and $L \equiv \sum_{i \in I} L_i$ respectively. We define the aggregate elasticity of substitution, σ^{agg}, to be the partial equilibrium response of the aggregate capital-labor ratio, K/L, to a change in relative factor prices, w/r:[11]

$$\sigma^{agg} \equiv \frac{d \ln K/L}{d \ln w/r} \tag{2}$$

We neither impose nor derive a parametric form for an aggregate production function. Given the allocation of capital and labor, σ^{agg} simply summarizes, to a first order, how a change in factor prices would affect the aggregate capital-labor ratio.

Let $\alpha_i \equiv \frac{rK_i}{rK_i + wL_i}$ and $\alpha \equiv \frac{rK}{rK + wL}$ denote the cost shares of capital for plant i and in aggregate. The plant-level and industry-level elasticities of substitution are closely related to the changes in these capital shares:

$$\begin{aligned} \sigma - 1 &= \frac{d \ln rK_i/wL_i}{d \ln w/r} = \frac{d \ln \alpha_i/(1-\alpha_i)}{d \ln w/r} = \frac{1}{\alpha_i(1-\alpha_i)} \frac{d\alpha_i}{d \ln w/r} && (3) \\ \sigma^{agg} - 1 &= \frac{d \ln rK/wL}{d \ln w/r} = \frac{d \ln \alpha/(1-\alpha)}{d \ln w/r} = \frac{1}{\alpha(1-\alpha)} \frac{d\alpha}{d \ln w/r} && (4) \end{aligned}$$

The aggregate cost share of capital can be expressed as an average of plant capital shares, weighted by size:

$$\alpha = \sum_{i \in I} \alpha_i \theta_i \tag{5}$$

where $\theta_i \equiv \frac{rK_i + wL_i}{rK + wL}$ denotes plant i's expenditure on capital and labor as a fraction of the aggregate expenditure. To find the aggregate elasticity of substitution, we can simply differentiate

[11] Since production and demand are homogenous of degree one, a change in total spending would not alter its capital-labor ratio. We address non-homothetic environments in Appendix B.6.

equation (5):

$$\frac{d\alpha}{d\ln w/r} = \sum_{i\in I}\frac{d\alpha_i}{d\ln w/r}\theta_i + \sum_{i\in I}\alpha_i\frac{d\theta_i}{d\ln w/r}$$

Using equations 3 and 4, this can be written as

$$\sigma^{agg}-1 = \frac{1}{\alpha(1-\alpha)}\sum_{i\in I}\alpha_i(1-\alpha_i)(\sigma-1)\theta_i + \frac{1}{\alpha(1-\alpha)}\sum_{i\in I}\alpha_i\theta_i\frac{d\ln\theta_i}{d\ln w/r}$$

The first term on the right hand side is a substitution effect that captures the change in factor intensity holding fixed each plant's size, θ_i. σ measures how much an individual plant changes its mix of capital and labor in response to changes in factor prices. The second term is a reallocation effect that captures how plants' size changes with relative factor prices. By Shephard's Lemma, a plant's cost share of capital α_i measures how relative factor prices affect its marginal cost. When wages rise, plants that use capital more intensively gain a relative cost advantage. Consumers respond to the subsequent changes in relative prices by shifting consumption toward the capital intensive goods. This reallocation effect is larger when demand is more elastic. More formally, the change in i's relative expenditure on capital and labor is

$$\frac{d\ln\theta_i}{d\ln w/r} = (\varepsilon-1)(\alpha_i-\alpha) \tag{6}$$

After some manipulation (see Appendix A for details), we can show that the industry elasticity of substitution is a convex combination of the micro elasticity of substitution and elasticity of demand:

$$\sigma^{agg} = (1-\chi)\sigma + \chi\varepsilon \tag{7}$$

where $\chi \equiv \sum_{i\in I}\frac{(\alpha_i-\alpha)^2}{(1-\alpha)\alpha}\theta_i$.

The first term, $(1-\chi)\sigma$, measures substitution between capital and labor within plants. The second term, $\chi\varepsilon$, captures reallocation between capital- and labor-intensive plants.

We call χ the heterogeneity index. It is proportional to the cost-weighted variance of capital shares and lies between zero and one.[12] When each plant produces at the same capital intensity, χ is zero and there is no reallocation across plants. Each plant's marginal cost responds to factor price changes in the same way, so relative output prices are unchanged. In contrast, if some plants produce using only capital while all others produce using only labor, all factor substitution is across plants and χ is one. When there is little variation in capital intensities, within-plant substitution is more important than reallocation.

2.2 Baseline Model

This section describes the baseline model we will use in our empirical implementation. The baseline model extends the previous analysis by allowing for heterogeneity across industries and using a production structure in which plants use materials in addition to capital and labor.

Let N be the set of industries and I_n be the set of plants in industry n. We assume that each plant's production function has a nested CES structure.

Assumption 1 *Plant i in industry n produces with the production function*

$$F_{ni}\left(K_{ni}, L_{ni}, M_{ni}\right) = \left(\left[\left(A_{ni}K_{ni}\right)^{\frac{\sigma_n-1}{\sigma_n}} + \left(B_{ni}L_{ni}\right)^{\frac{\sigma_n-1}{\sigma_n}}\right]^{\frac{\sigma_n}{\sigma_n-1}\frac{\zeta_n-1}{\zeta_n}} + \left(C_{ni}M_{ni}\right)^{\frac{\zeta_n-1}{\zeta_n}}\right)^{\frac{\zeta_n}{\zeta_n-1}} \tag{8}$$

so its elasticity of substitution between capital and labor is σ_n. Further, i's elasticity of substitution between materials and its capital-labor bundle is ζ_n.

We also assume that demand has a nested structure with a constant elasticity at each level of

[12] A simple proof: $\sum_{i\in I}(\alpha_i-\alpha)^2\theta_i = \sum_{i\in I}\alpha_i^2\theta_i - \alpha^2 \le \sum_{i\in I}\alpha_i\theta_i - \alpha^2 = \alpha - \alpha^2 = \alpha(1-\alpha)$. It follows that $\chi = 1$ if and only if each plant uses only capital or only labor (i.e., for each i, $\alpha_i \in \{0,1\}$).

aggregation. Such a structure is consistent with a representative consumer whose preferences exhibit constant elasticities of substitution across industries and across varieties within each industry:

$$Y \equiv \left[\sum_{n \in N} D_n^{\frac{1}{\eta}} Y_n^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}, \qquad Y_n \equiv \left(\sum_{i \in I_n} D_{ni}^{\frac{1}{\varepsilon_n}} Y_{ni}^{\frac{\varepsilon_n-1}{\varepsilon_n}}\right)^{\frac{\varepsilon_n}{\varepsilon_n-1}} \tag{9}$$

This demand structure implies that each plant in industry n faces a demand curve with constant elasticity ε_n. Letting q be the price of materials, each plant maximizes profit

$$\max_{P_{ni}, Y_{ni}, K_{ni}, L_{ni}, M_{ni}} P_{ni} Y_{ni} - rK_{ni} - wL_{ni} - qM_{ni}$$

subject to the technological constraint $Y_{ni} = F_{ni}(K_{ni}, L_{ni}, M_{ni})$ and the demand curve $Y_{ni} = Y_n (P_{ni}/P_n)^{-\varepsilon_n}$, where $P_n \equiv \left(\sum_{i \in I_n} D_{ni} P_{ni}^{1-\varepsilon_n}\right)^{\frac{1}{1-\varepsilon_n}}$ is the price index for industry n.

The industry-level elasticity of substitution between capital and labor for industry n measures the response of the industry's capital-labor ratio to a change in relative factor prices:

$$\sigma_n^N \equiv \frac{d \ln K_n / L_n}{d \ln w/r}$$

The derivation of this industry elasticity of substitution follows Section 2.1 up to equation (6). As before, $\alpha_{ni} = \frac{rK_{ni}}{rK_{ni}+wL_{ni}}$ is plant i's capital share of non-materials cost and $\theta_{ni} = \frac{rK_{ni}+wL_{ni}}{rK_n+wL_n}$ plant i's share of industry n's expenditure on capital and labor. We will show that reallocation depends on plants' expenditures on materials. We denote plant i's materials share of its total cost as $s_{ni}^M \equiv \frac{qM_{ni}}{rK_{ni}+wL_{ni}+qM_{ni}}$. Because producers of intermediate inputs use capital and labor, changes in r and w would impact the price of materials. We define $\alpha^M \equiv \frac{d \ln q/w}{d \ln r/w}$ to be the capital content of materials.

Proposition 1 *Under Assumption 1, the industry elasticity of substitution is:*

$$\sigma_n^N = (1-\chi_n)\sigma_n + \chi_n\left[(1-\bar{s}_n^M)\varepsilon_n + \bar{s}_n^M\zeta_n\right]$$

where $\chi_n = \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)^2}{(1-\alpha_n)\alpha_n}\theta_{ni}$ *and* $\bar{s}_n^M = \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)(\alpha_{ni}-\alpha^M)\theta_{ni}s_{ni}^M}{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)(\alpha_{ni}-\alpha^M)\theta_{ni}}$

The proofs of all propositions are contained in Appendix A.

Relative to equation (7), the demand elasticity is replaced by a convex combination of the elasticity of demand, ε_n, and the elasticity of substitution between materials and the capital-labor bundle, ζ_n. This composite term measures the change in i's share of its industry's expenditure on capital and labor, θ_{ni}. Intuitively, a plant's expenditure on capital and labor can fall because its overall scale declines or because it substitutes towards materials. The cost share of materials determines the relative importance of each. As materials shares approach zero, all shifts in composition are due to changes in scale, and Proposition 1 reduces to equation (7). In contrast, as a plant's materials share approaches one, changes in its cost of capital and labor have a negligible impact on its marginal cost, and hence a negligible impact on its sales. Rather, the change in its expenditure on capital and labor is determined by substitution between materials and the capital-labor bundle.

The aggregate elasticity parallels the industry elasticity; aggregate capital-labor substitution consists of substitution within industries and reallocation across industries. Proposition 2 shows that expression for the aggregate elasticity parallels the expressions for the industry elasticity in Proposition 1 with plant and industry variables replaced by industry and aggregate variables respectively.

Proposition 2 *The aggregate elasticity between capital and labor,* $\sigma^{agg} = \frac{d\ln K/L}{d\ln w/r}$*, is:*

$$\sigma^{agg} = (1-\chi^{agg})\,\bar{\sigma}^N + \chi^{agg}\left[(1-\bar{s}^M)\eta + \bar{s}^M\bar{\zeta}^N\right] \tag{10}$$

where $\chi^{agg} \equiv \sum_{n\in N} \frac{(\alpha_n-\alpha)^2}{\alpha(1-\alpha)}\theta_n$, $\bar{s}^M \equiv \sum_{n\in N} \frac{(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n}{\sum_{n'\in N}(\alpha_{n'}-\alpha)(\alpha_{n'}-\alpha^M)\theta_{n'}} s_n^M$, $\bar{\sigma}^N \equiv \sum_{n\in N} \frac{\alpha_n(1-\alpha_n)\theta_n}{\sum_{n'\in N}\alpha_{n'}(1-\alpha_{n'})\theta_n}\sigma_n^N$, *and* $\bar{\zeta}^N \equiv \sum_{n\in N} \frac{(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n s_n^M}{\sum_{n'\in N}(\alpha_{n'}-\alpha)(\alpha_{n'}-\alpha^M)\theta_{n'} s_{n'}^M}\zeta_n^N$.

Substitution within industries depends on $\bar{\sigma}^N$, a weighted average of the industry elasticities of substitution defined in Proposition 1. $\bar{\zeta}^N$ is similarly a weighted average of industry level elasticities of substitution between materials and non-materials (we relegate the expression ζ_n^N to Appendix A).

3 The US Aggregate Elasticity of Substitution

The methodology developed in the previous section shows how to recover the aggregate capital-labor elasticity from micro parameters and the distribution of plant expenditures. We now use plant-level data on US manufacturing plants to estimate all of the micro components of the aggregate elasticity. We then assemble these components to estimate the aggregate capital-labor elasticity of substitution for the US manufacturing sector and examine its behavior over time.

3.1 Data

Our two main sources of micro data on manufacturing plants are the US Census of Manufactures and Annual Survey of Manufactures (ASM). The Census of Manufactures is a census of all manufacturing plants conducted every five years. It contains more than 180,000 plants per year.[13] The Annual Survey of Manufactures tracks about 50,000 plants over five year panel rotations between Census years, and includes the largest plants with certainty.

In this study, we primarily use factor shares measured at the plant level. For the Census samples, we measure capital by the end year book value of capital, deflated using an industry specific current cost to historic cost deflator. The ASM has the capital and investment history

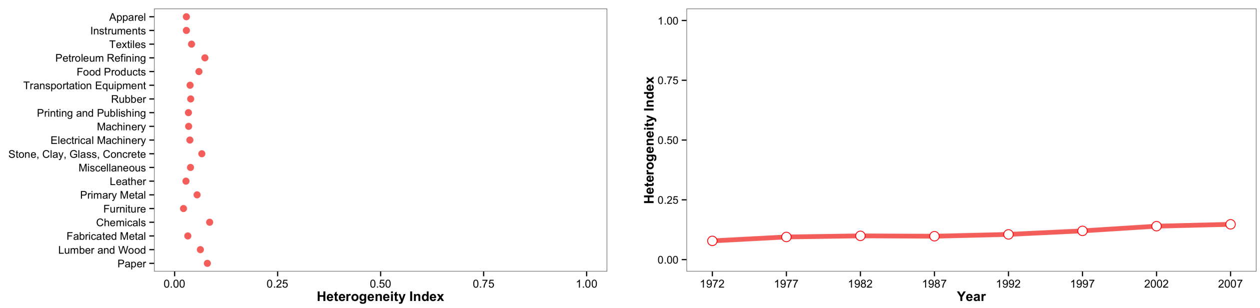
[13]This excludes small Administrative Record plants with fewer than five employees, for whom the Census only tracks payroll and employment. We omit these in line with the rest of the literature using manufacturing Census data.

required to construct perpetual inventory measures of capital. We thus create perpetual inventory measures of capital, accounting for retirement data when possible as in Caballero et al. (1995) and using NIPA investment deflators for each industry-capital type. Capital costs consist of the total stock of structures and equipment capital multiplied by the appropriate rental rate, using rental rates based upon an external real rate of return of 3.5 percent as in Harper et al. (1989) and tax rates from Dale Jorgenson. For labor costs, both surveys contain total salaries and wages at the plant level, but the ASM subsample is supplemented with data on supplemental labor costs including benefits as well as payroll and other taxes. For details about data construction, see Appendix C.

The Census of Manufactures, unlike the ASM subsample, only contains capital data beginning in 1987. Further, industry definitions change from SIC to NAICS in 1997. Given all of these considerations, we take the following approach to estimating the aggregate elasticity.

Our approach is to use the full 1987 Census of Manufactures to estimate the micro elasticities and examine their robustness. We then use the ASM in each year for the relevant information on the composition of plants- the heterogeneity indices and materials shares- because we extend the analysis from 1972 to 2007. So, for example, to compute the aggregate elasticity of substitution in 1977, we combine estimates of micro parameters from the 1987 Census of Manufactures with information from the 1977 ASM.

Throughout, we use a plant's total cost of labor as a measure of its labor input. We view employees of different skill as supplying different quantities of efficiency units of labor, so using the wage bill controls for differences in skill. In Section 4.3 we show that our methodology is valid even if wages per efficiency unit of labor vary across plants.



(a) Heterogeneity Indices Across Industries

(b) Heterogeneity Indices over Time

Figure 1 Heterogeneity Indices

Note: The left figure displays the heterogeneity index, χ_n, in 1987 for each industry. The right figure displays the average heterogeneity index over time.

3.2 Micro Heterogeneity

The extent of heterogeneity in capital intensities, as measured by the heterogeneity index, determines the relative importance of within-plant substitution and reallocation. Figure 1a depicts these indices for each industry in 1987. Across industries, the indices average 0.1 and are all less than 0.2. Similarly, Figure 1b shows how the average heterogeneity index evolves over time. While heterogeneity indices are rising, they remain relatively small. Industry capital shares exhibit even less variation ($\chi^N = 0.05$).

Given this level of heterogeneity, the plant-level elasticity of substitution between capital and labor is a primary determinant of the aggregate elasticity. We therefore begin with a thorough investigation of this elasticity.

3.3 Plant Level Elasticity of Substitution

We obtain the plant-level elasticity of substitution from Raval (2013). We describe the methodology and estimates in detail in order to explain how we map the theory to the data, and then compare these estimates to others from the literature.

Given cost minimization, the relationship between relative factor costs of capital and labor

rK_i/wL_i and relative factor prices w/r identifies this elasticity. We exploit wage differences across local areas in the US in order to identify the micro elasticity of substitution between capital and labor. Because these wage differences are persistent, they identify plants long-run response to a permanent change in factor prices. We run the regression:

$$\log \frac{rK_i}{wL_i} = (\sigma_n - 1) \log w_i + CONTROLS + \epsilon_i$$

where i denotes an individual plant and w_i the wage for the MSA in which the plant is located. The implicit assumption is that capital is mobile so all plants face the same cost of capital. To obtain a MSA level wage, we first estimate a residual wage for each person after controlling for education, experience, and demographics using data from the Population Censuses. We then average this residual within the MSA.[14] All regressions control for industry fixed effects, as well as plant age and multi-unit status.[15]

This specification has several attractive properties. First, the dependent variable uses a plant's wage bill rather than a count of employees. If employees supply efficiency units of labor, using the wage bill automatically controls for differences in skill across workers and, more importantly, across plants. Second, plants may find it costly to adjust capital or labor. Deviations of a plant's capital or labor from static cost minimization due to adjustment costs would be in the residual, but are orthogonal to the MSA wage. Third, the MSA wage and plant wage bills are calculated using different data sources, so we avoid division bias from measurement error in the wage in the dependent and independent variables.

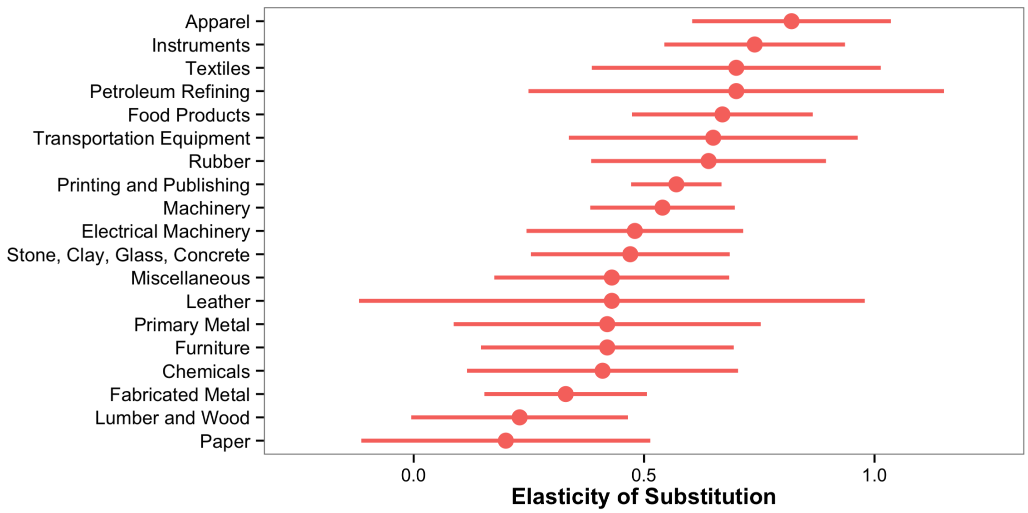
Using all manufacturing plants, Raval (2013) estimates a plant level elasticity of substitution close to one-half in both 1987 and 1997. In this paper, we allow plant elasticities of substitution to

[14] For details about how we construct this wage, see Appendix C.2.

[15] The regression only uses plants in a single year. The implicit assumption is that capital is mobile so the rental rate of capital is the same in all MSAs. We examine this assumption below.

vary by two digit SIC industry and so run separate regressions for each industry. Figure 2 displays the estimates by industry along with the 95 percent confidence interval. Most of the estimates range between 0.4 and 0.7. The remainder of this section addresses a number of potential issues with our plant level elasticity estimates.

Figure 2 Plant Elasticity of Substitution by Industry, 1987



Note: For each industry, this graph plots the plant level elasticity of substitution between capital and labor as estimated in Raval (2013), together with the 95 percent confidence interval for each estimate. Standard errors are clustered at the MSA-industry level.

Endogeneity

We estimated the plant level elasticity of substitution using cross-sectional wage differences across US locations. A natural question is whether these wage differences are exogenous to non-neutral productivity differences. For example, to the extent that higher wages in MSAs were caused by higher labor-augmenting productivity or unobserved skills, our estimate of the elasticity of substitution would be biased towards one as $\sigma - 1$ would be attenuated.

To address such endogeneity problems, we use a version of Bartik (1991)'s instrument for labor demand, which is based on the premise that MSAs differ in their industrial composition. When an industry expands nationwide, MSAs more heavily exposed to that industry experience larger increases in labor demand. Thus given each MSA's initial industrial composition, we can construct the change in each MSA's labor demand due to the change in each industry's nationwide employment. We restrict the instrument to non-manufacturing industries only.[16]

Table I contains estimates of the elasticity of substitution using these instruments in the third column.[17] The first two columns contain our baseline least squares estimates. The first column contains the average plant-level elasticity when we estimate separate elasticities for each two-digit SIC industry. The second column estimates a common elasticity across all industries in manufacturing. In both years, the IV estimates are close to one-half, ranging from 0.49 to 0.52, and are similar to the baseline least squares estimates.

Table I Robustness Checks for Plant Capital-Labor Substitution Elasticity

	Separate OLS	Single OLS	Bartik Instrument	Equipment Capital	Firm FE
1987	0.52	0.52 *(0.04)*	0.49 *(0.05)*	0.53 *(0.03)*	0.49 *(0.05)*
1997	0.52	0.46 *(0.03)*	0.52 *(0.08)*		0.48 *(0.08)*

Note: Standard errors are in parentheses. The table contains five specifications. The first specification estimates a separate plant elasticity of substitution for each industry, and then averages them using the cross industry weights used for aggregation. The second specification estimates a single common elasticity of substitution for the entire manufacturing sector. The third specification estimates the elasticity of substitution in an IV framework using Bartik labor demand instruments. The fourth specification estimates the plant elasticity of substitution through least squares, defining capital to only include equipment capital. The fifth specification estimates the plant elasticity of substitution through least squares but includes firm fixed effects, excluding all single plant firms.

All regressions include industry fixed effects, age fixed effects, and a multi-unit status indicator. Wages used are the average log wage for the MSA. In the IV specification, the wage is computed as payroll/number of employees at the establishment level; in all other cases, the wage is computed as wage and salary income over total number of hours worked adjusted for differences in worker characteristics. Standard errors are clustered at the industry-MSA level.

Rental Rate of Capital

Our estimate of the micro elasticity of substitution would be biased if rental rates vary sys-

[16] Formally, the instrument is constructed as follows: Let $g_n(t) = \frac{1}{10}\ln(L_n(t)/L_n(t-10))$ be the national growth rate of industry n, and let $\omega_{j,n}(t)$ be the share of MSA j's employment working in industry j. Finally, the instrument is the interaction between initial MSA employment shares and 10 year national employment growth rates: $Z_j(t) = \sum_{n\in N^S}\omega_{j,n}(t-10)g_n(t)$, where N^S is the set of non-manufacturing four-digit SIC industries.

[17] While all other specifications use wages from data on workers from the Population Censuses, here we use wages based on data on average payroll to employment across all establishments in an MSA from the Longitudinal Business Database. The Population Censuses are only conducted every 10 years in different years from the Economic Censuses, so the wages from the Population Censuses do not match the year of the Economic Census. For most of our specifications, this mismatch is not a problem because our wage variation is highly persistent. This mismatch becomes a problem if we want to use short run variation in wages from labor demand shocks. While the wages from establishment data do not control for differences in individual worker characteristics, the labor demand instrument should be orthogonal to the measurement error in wages. Because the SIC industry definitions changed from 1972 SIC basis to 1987 SIC basis in 1987, for 1987 we use the 1976-1986 instrument.

tematically with local wages. There are two reasons that might cause these rental rates to vary with wages. First, the cost of some kinds of capital, such as structures, may reflect local wages. To examine this, we estimate the elasticity of substitution between labor and equipment capital, which is more plausibly mobile across locations.[18] Second, the cost of capital could vary through differences in lending rates from banks in different locations, or from differences in firm creditworthiness or access to capital markets. To control for these differences, we add firm fixed effects for plants that belong to multi-unit firms. The fourth column contains estimates from the equipment capital specification while the fifth column in Table I contains estimates from the fixed effect specification; in both cases, the estimates are quite close to those from the baseline specification.

Extensive Margin

In the baseline model presented in Section 2, plants could substitute across inputs and change size. However, the set of existing plants remained fixed, so there was no extensive margin of adjustment. Would a higher wage cause entering firms to choose more capital intensive technologies? If so, the aggregate elasticity of substitution should account for that shift.

Consider a putty-clay model in which plant i has some core characteristics $\{\mathcal{A}_i, \mathcal{B}_i\}$. Upon entering, it can choose a technology $\{A_i, B_i\}$ from the menu

$$\left[(A_i/\mathcal{A}_i)^{1-\sigma^{ext}} + (B_i/\mathcal{B}_i)^{1-\sigma^{ext}}\right]^{\frac{1}{1-\sigma^{ext}}} \leq 1$$

Once it has selected its technology, it cannot change to an alternative technology. After entering, it produces using the production function

$$Y_i = \left[(A_iK_i)^{\frac{\sigma^{int}-1}{\sigma^{int}}} + (B_iL_i)^{\frac{\sigma^{int}-1}{\sigma^{int}}}\right]^{\frac{\sigma^{int}}{\sigma^{int}-1}} \tag{11}$$

[18] Because the Census does not separate equipment and structures capital in 1997, we only estimate this specification for 1987.

σ^{int} thus represents a short run elasticity. Once the plant has entered, it cannot switch $\{A_i, B_i\}$, so σ^{int} is the response of i's capital-labor ratio to relative factor prices:

$$\frac{K_i}{L_i} = \left(\frac{A_i}{B_i}\right)^{\sigma^{int}-1} \left(\frac{r}{w}\right)^{-\sigma^{int}}$$

A shift in factor prices would also alter entering plants' choices of technologies. Given factor prices, i's choice of technologies will satisfy $\left(\frac{A_i K_i}{B_i L_i}\right)^{\frac{\sigma^{int}-1}{\sigma^{int}}} = \left(\frac{A_i/\mathcal{A}_i}{B_i/\mathcal{B}_i}\right)^{\sigma^{ext}-1}$. Along with i's choice of capital and labor, this implies that, after entry, the entering plant's capital-labor ratio will be

$$\frac{K_i}{L_i} = \left(\frac{\mathcal{A}_i}{\mathcal{B}_i}\right)^{\sigma^{total}-1} \left(\frac{r}{w}\right)^{-\sigma^{total}}$$

where σ^{total} is defined to satisfy

$$\frac{1}{\sigma^{total}-1} = \frac{1}{\sigma^{int}-1} + \frac{1}{\sigma^{ext}-1}$$

σ^{total} represents a long run elasticity of substitution. If, for example, the wage was high and remained high, all entering plants would choose more capital intensive technologies, and σ^{total} would capture the resulting shifts in capital-labor ratios.

If the true model contains both an intensive and extensive margin, how should we interpret our estimates of the micro elasticity? Our estimation strategy uses cross-sectional differences; we compare capital-labor ratios across locations with different wages. These differences in capital-labor ratios come from some combination of the intensive and extensive margins. Because geographic wage differences are extremely persistent, our estimated micro elasticity corresponds to σ^{total}; in high-wage locations, past entering cohorts would have selected technologies that reflected the higher cost of labor.

With our methodology we are unable to distinguish between the intensive and extensive margins of adjustment. Fortunately, doing so is not necessary to build up to a long-run aggregate elasticity.[19]

Sorting

Our estimates do not account for the possibility that plants select locations in response to factor prices. To see why this might matter, consider the following extreme example: Suppose plants cannot adjust their factor usage but can move freely. Then we would expect to find more labor intensive plants in locations with lower wages. However, a national increase in the wage would not change any plant's factor usage. Thus to the extent that this channel is important, our estimated elasticity will overstate the true elasticity.

Plants ability to sort across locations likely varies by industry. We would expect industries in which plants are more mobile to be more clustered in particular areas. This could depend, for example, on how easily goods can be shipped to other locations.

Raval (2013) addresses this by looking at a set of ten large four-digit industries located in almost all MSAs and states. These are industries for which we would expect sorting across locations to be least important. The leading example of this is ready-mixed concrete; because concrete cannot be shipped very far, concrete plants exist in every locality. Elasticities for these industries are similar to the estimates for all industries in our baseline, with average elasticities of 0.49 for 1987 and 0.61 for 1997.

Alternative Estimates

Our estimation strategy uses cross-sectional variation. Chirinko (2008) surveys the existing

[19] Houthakker (1955) also features an extensive margin. The argument that our estimate captures both the intensive and extensive margins of adjustment is the same, but the mapping from Houthakker's model to our parameter estimates is more opaque. In that model, even though individual plants have Leontief production functions, one can show that the equilibrium distribution of capital shares (α_i) in the cross section is independent of factor prices. Thus we would estimate a unit plant-level elasticity of substitution. Houthakker also assumed that each plant's capacity was constrained by a fixed factor, so changes in factor prices do not alter plants' scales. This corresponds to a unit demand elasticity in our baseline model. Thus, for any heterogeneity index, our estimation strategy would imply a Cobb-Douglas aggregate production function.

literature using micro data and time series variation. This literature typically uses variation in the user cost of capital over time to identify the elasticity; some of this variation stems from tax-law changes that differentially affect different asset types, and so different industries. Chirinko (2008), summarizing Chirinko et al. (2011), argues that the weight of the evidence indicates a long-run micro capital-labor elasticity of substitution between 0.4 and 0.6. We find it comforting that two approaches that use very different sources of variation yield similar estimates.

3.4 Aggregation

We now estimate the remaining plant-level production and demand parameters. We then use them to aggregate to the manufacturing-level elasticity of substitution.

The reallocation effect depends upon both the plant elasticity of substitution between materials and non-materials, ζ, and the elasticity of demand, ε.

To identify ζ, we use the same cross-area variation in the wage. Across MSAs, the local wage varies but the prices of capital and materials are fixed. Cost minimization implies that ζ measures the response of relative expenditures of materials and non-materials to changes in their respective prices: $1-\zeta = \frac{d\ln[(rK_i+wL_i)/qM_i]}{d\ln(\lambda_i/q)}$, where λ_i is the cost index of i's capital-labor bundle. Holding fixed the prices of materials and capital, a change in the local wage would increase these relative prices by $d\ln\lambda_i/q = (1-\alpha_i)d\ln w_i$. Thus to a first-order approximation, the response of $(rK_i+wL_i)/qM_i$ to the local wage would thus be $(1-\zeta)(1-\alpha_i)$. We therefore estimate ζ using the regression

$$\log\frac{rK_i + wL_i}{qM_i} = (1-\zeta)(1-\alpha_i)(\log w_i) + CONTROLS + \epsilon_i$$

Table II contains these elasticities for 1987 and 1997. Because we use the full Census for each estimate, our estimate of ζ is common across industries. The first column contains our baseline estimate. This implicitly assumes a national market for materials. The second column adjusts the regression to account for the fact that some materials are sourced locally; an increase in the local wage would raise the cost of such locally sourced materials. See Appendix D.1 for details. These estimates are a bit lower than one; in our subsequent calculations, we use the 1987 MSA level estimate of 0.90.[20]

Table II Elasticities of Substitution between Materials and Non-Materials for the Manufacturing Sector

	No Local Content	Local Content
1987	0.90 *(0.06)*	0.87 *(0.07)*
1997	0.67 *(0.04)*	0.63 *(0.05)*
N	$\approx 140,000$	

Note: Standard errors are in parentheses. All regressions include industry fixed effects, age fixed effects and a multi-unit status indicator, and have standard errors clustered at the two digit industry-MSA level. The wages used are the average log wage for the MSA, where the wage is computed as wage and salary income over total number of hours worked adjusted for differences in education, experience, race, occupation, and industry. Specifications with local content are as described in Appendix D.1.

Within industries, the demand elasticity tells us how much consumers substitute across plants when relative prices change. We estimate the elasticity of demand using the implications of profit maximization; optimal price setting behavior implies that the markup over marginal cost is equal to $\frac{\varepsilon}{\varepsilon-1}$. We thus invert the average markup across plants in an industry to obtain the elasticity of demand. The assumption of constant returns to scale implies that each plant's markup is equal to its ratio of revenue to cost.[21] Figure 3 displays the elasticities of demand for each manufacturing industry in 1987. Across industries, elasticities of demand vary between three and five.

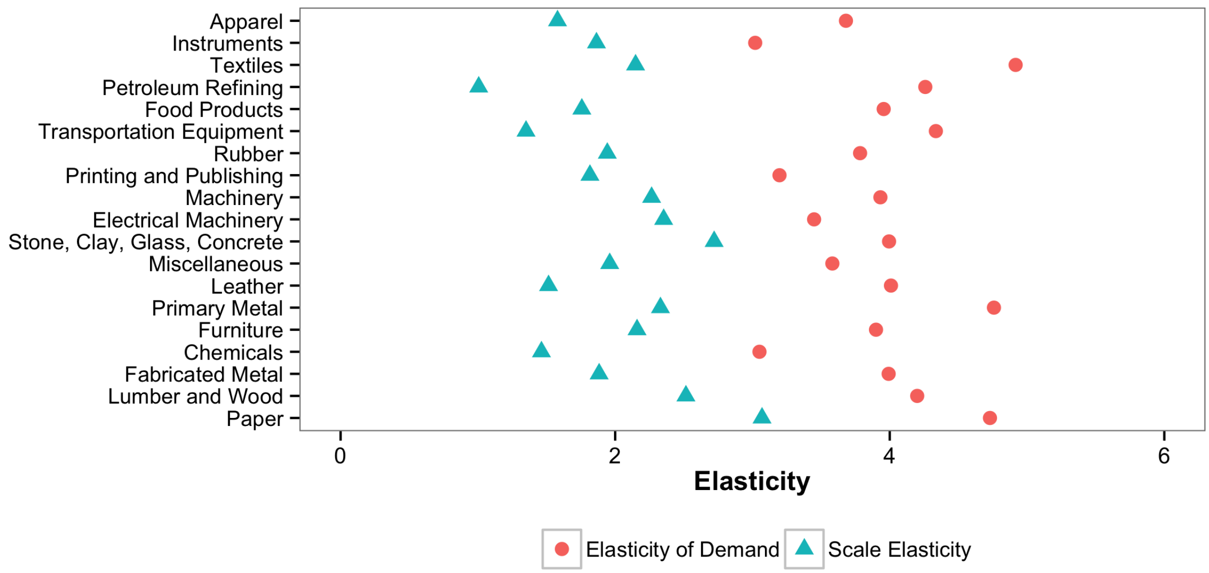
The overall scale elasticity is $\bar{s}_n^M \zeta + (1 - \bar{s}_n^M)\varepsilon_n$. $\bar{s}_n^M$ is an average of materials shares, which are high in manufacturing; the average across industries in 1987 is 0.59. Figure 3 contains our estimates of the scale elasticities; they average 2.12 across industries.[22]

[20] Atalay (2014) pursues a complementary approach using differences in materials prices across plants in the US Census of Manufactures and finds estimates within the range of Table II. This differs from his main estimate of this elasticity which uses shorter-run industry-level variation, and so may not reflect the long-run, plant-level elasticity required for our model.

[21] We relax the assumption of constant returns to scale in Section 4.2, and explore alternative estimation strategies for the demand elasticity in Section 4.1.

[22] This average, along with other averages across industries, is a weighted average where the weight on industry n is $\frac{\alpha_n(1-\alpha_n)\theta_n}{\sum_{n\in N}\alpha_n(1-\alpha_n)\theta_n}$ as in Proposition 2.

Figure 3 Elasticity of Demand and Scale Elasticity by Industry, 1987

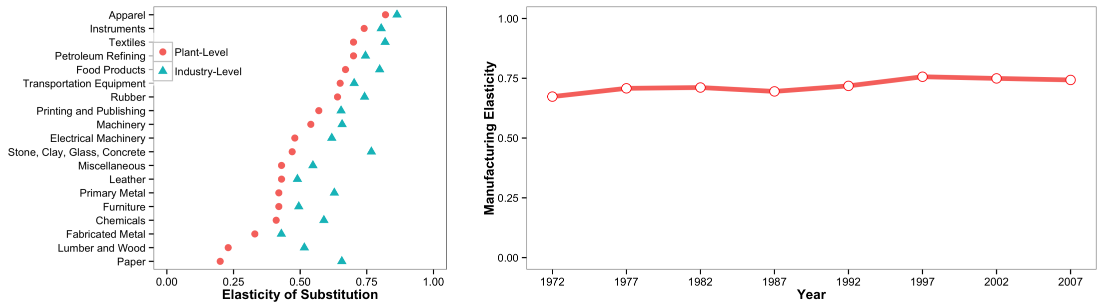


Note: For each industry, this graph plots both the elasticity of demand, estimated from revenue-cost ratios, and the scale elasticity $\bar{s}_n^M \zeta + (1 - \bar{s}_n^M)\varepsilon_n$.

To aggregate across industries, we need one more elasticity, the cross industry elasticity of demand η. We estimate this elasticity using industry-level panel data by regressing quantity on price, using average cost as an instrument for supply. Appendix D.2 contains the details of this analysis. Across specifications, we find estimates centered around one. We thus set η to one. As we would expect, the cross-industry demand elasticity η is much lower than the plant-level demand elasticities; varieties in the same industries are better substitutes than varieties in other industries.

We can now combine the substitution and reallocation effects to estimate the industry and manufacturing sector level elasticities of substitution. In Figure 4a, we depict the plant level and industry level elasticities of substitution. Because the heterogeneity indices tend to be small, the industry-level elasticities of substitution are only moderately higher than the plant-level elasticities. The average industry elasticity is 0.66 and the overall manufacturing level elasticity of substitution is 0.70.[23] Within-plant substitution accounts for 71 percent of industry substitution and 65 percent

[23]For our methodology, uncertainty about specification is likely more important than statistical imprecision. That being said, the standard error of our estimate of the aggregate capital-labor elasticity is 0.03. We arrive at this by



(a) Plant and Industry Elasticities

(b) Aggregate Elasticity Across Time

Figure 4 Elasticities of Substitution and Aggregation

Note: The left figure displays the plant level elasticity and industry level elasticity of substitution for each industry. The right figure displays the manufacturing level elasticity of substitution for each Census year from 1972-2007.

of overall substitution for manufacturing.

Our methodology now allows us to examine the stability of the aggregate elasticity of substitution over time by allowing the heterogeneity across plants and industry composition to vary across time.[24] Figure 4b depicts the aggregate elasticity of substitution from 1972 to 2007. The aggregate elasticity has risen slightly from 0.67 to 0.73. In Appendix D.3, we examine the robustness of this approach by using estimates of the micro production and demand elasticities based on the 1997 Census of Manufactures. The resulting time path of aggregate elasticities is virtually identical to those that use the 1987 micro estimates.

4 Robustness

In this section, we examine the robustness of our estimates of the micro elasticities of demand. We also discuss changes in the specification of the economic environment, including imperfect pass-through, returns to scale, adjustment costs, and other misallocation frictions.

using the standard errors from our micro level elasticity estimates.

[24] After 1997, industry definitions switch from two digit SIC basis to three digit NAICS basis. We assign SIC plant elasticities to the equivalent NAICS industries.

4.1 Elasticity of Demand

In this section, we examine the robustness of our estimates of the elasticity of demand. For several homogenous products, the US Census of Manufactures collects both price and physical quantity data. We can thus estimate the elasticity of demand by regressing quantity on price, instrumenting for price using average cost. This approach is similar to Foster et al. (2008). This strategy implies an average industry-level capital-labor elasticity of substitution of 0.52 among these industries, close to the estimate of 0.54 using our baseline strategy.[25] The trade literature finds estimates in the same range using within industry variation across imported varieties to identify the elasticity of demand. For example, Imbs and Mejean (2011) find a median elasticity of 4.1 across manufacturing industries.

Appendix B.5 generalizes the analysis to a homothetic demand system with arbitrary demand elasticities and imperfect pass-through of marginal cost. In that environment, the formula for the industry elasticity in Proposition 1 is unchanged except the elasticity of demand ε_n is replaced by a weighted average of the quantity $b_i\varepsilon_i$; ε_i is i's local demand elasticity and b_i is i's local rate of relative pass-through (the elasticity of its price to a change in marginal cost). Note that under Dixit-Stiglitz preferences, $b_i = 1$ and $\varepsilon_i = \varepsilon_n$ for each i. Here, however, if a plant passes through only half of a marginal cost increase, then the subsequent change in scale would be half as large.

4.2 Returns to Scale

Our baseline estimation assumed that each plant produced using a production function with constant returns to scale. Alternatively, we can assume that plant i produces using the production

[25] Foster et al. (2008) instrument for price using plant-level TFP. We cannot use their estimates directly because they assume plants produce using homogeneous Cobb-Douglas production functions. Because we maintain the assumption of constant returns to scale, the appropriate analogue to plant-level TFP is average cost. Directly using the demand elasticities of Foster et al. (2008) would yield an average industry-level elasticity of substitution of 0.54.

function:

$$Y_i = F_i(K_i, L_i, M_i) = G_i(K_i, L_i, M_i)^\gamma$$

where G_i has constant returns to scale and $\gamma < \frac{\varepsilon}{\varepsilon - 1}$. Relative to the baseline, two things change, as shown in Appendix B.3. First, the industry elasticity of substitution becomes

$$\sigma_n^N = (1 - \chi_n)\bar{\sigma}_n + \chi_n \left[\bar{s}_n^M \bar{\zeta}_n + (1 - \bar{s}_n^M) x_n\right]$$

where x_n is defined to satisfy $\frac{x_n}{x_n - 1} = \frac{1}{\gamma}\frac{\varepsilon_n}{\varepsilon_n - 1}$. Thus the scale elasticity is a composite of two parameters, the elasticity of demand and the returns to scale. When the wage falls, the amount a labor-intensive plant would expand depends on both.

Second, when we divide a plant's revenue by total cost, we no longer recover the markup. Instead, we get

$$\frac{P_i Y_i}{rK_i + wL_i + qM_i} = \frac{1}{\gamma}\frac{\varepsilon_n}{\varepsilon_n - 1} = \frac{x_n}{x_n - 1}$$

Fortunately, this means that the procedure we used in the baseline delivers the correct aggregate elasticity of substitution even if we mis-specify the returns to scale. To see this, when we assumed constant returns to scale, we found the elasticity of demand by computing $\frac{\frac{P_i Y_i}{rK_i + wL_i + qM_i}}{\frac{P_i Y_i}{rK_i + wL_i + qM_i} - 1}$. With alternative returns to scale, this would no longer give the elasticity of demand, ε_n; rather, it gives the correct scale elasticity, x_n.[26]

4.3 Adjustment Costs and Distortions

Section 2 showed that the relative importance of within-plant substitution and reallocation depends upon the variation in cost shares of capital. Implicit in that environment was that this variation

[26] This specification of returns to scale rules out some features such as fixed costs of production. In an environment in which production functions exhibited such features, x_n would be replaced by a weighted average of individual plants' local scale elasticities. For details see Appendix B.6, which also derives an aggregate elasticity when production functions are non-homothetic.

came from non-neutral differences in technology. On the other hand, some of this heterogeneity may be due to adjustment costs or other distortions as in the recent misallocation literature (see Banerjee and Duflo (2005), Restuccia and Rogerson (2008), Hsieh and Klenow (2009)). A natural question arises: What are the implications for the aggregate elasticity of substitution if these differences come from idiosyncratic distortions?

Consider an alternative environment in which plants pay idiosyncratic prices for their inputs.[27] We are interested in how changes in factor prices impact the relative compensation of capital and labor. To be precise, suppose plant i pays factor prices $r_i = T_{Ki}r$ and $w_i = T_{Li}w$, and we define the industry elasticity of substitution in relation to the change in relative factor compensation in response to a change in w/r (holding fixed the idiosyncratic components of factor prices, $\{T_{Ki}, T_{Li}\}_{i\in I_n}$) so that it satisfies[28]

$$\sigma_n^N - 1 = \frac{d\ln\left(\sum_{i\in I_n} r_i K_i / \sum_{i\in I_n} w_i L_i\right)}{d\ln w/r}$$

As shown in Appendix B.4, it turns out that the expression for the industry elasticity of substitution is exactly the same as in Proposition 1, provided that we define $\alpha_i = \frac{r_i K_i}{r_i K_i + w_i L_i}$ and $\theta_i = \frac{r_i K_i + w_i L_i}{\sum_{i'\in I_n} r_{i'} K_{i'} + w_{i'} L_{i'}}$. Thus, as long as expenditures are measured correctly, no modifications are necessary.

Alternatively, a plant's shadow value of an input may differ from its marginal expenditure on that input. This could happen if the input is fixed in the short run or if use of that input is constrained by something other than prices. In the presence of such "unpaid" wedges, our expression for the industry elasticity would change slightly.

[27] In fact, our identification of the plant-level elasticity of substitution relies on plants (in different locations) facing different wages.

[28] In this environment the mapping between changes in the capital-labor ratio and changes in factor compensation is fuzzy; generically, $\frac{d\ln K_n/L_n}{d\ln w/r} - 1 \neq \frac{d\ln \sum_{i\in I_n} r_i K_i / \sum_{i\in I_n} w_i L_i}{d\ln w/r}$.

If we observed these unpaid wedges, we could compute the industry elasticity. However, measuring these wedges presents a challenge. Differences in plants' cost shares of capital could reflect differences in unpaid wedges or differences in technology; data on revenue and input expenditures alone are not sufficient to distinguish between the two.

To get a sense of how big of an issue these unpaid wedges might represent, we consider the following thought experiment. Suppose all variation in cost shares of capital were due to "unpaid" wedges. In that case, the aggregate elasticity of substitution for the manufacturing sector in 1987 would be 0.70, the same as our baseline, while in 1997 the aggregate elasticity would be 0.93 rather than 0.77 in the baseline (see Appendix B.4 for details). This suggests that misallocation would not substantially alter our analysis.

5 Bias of Technical Change

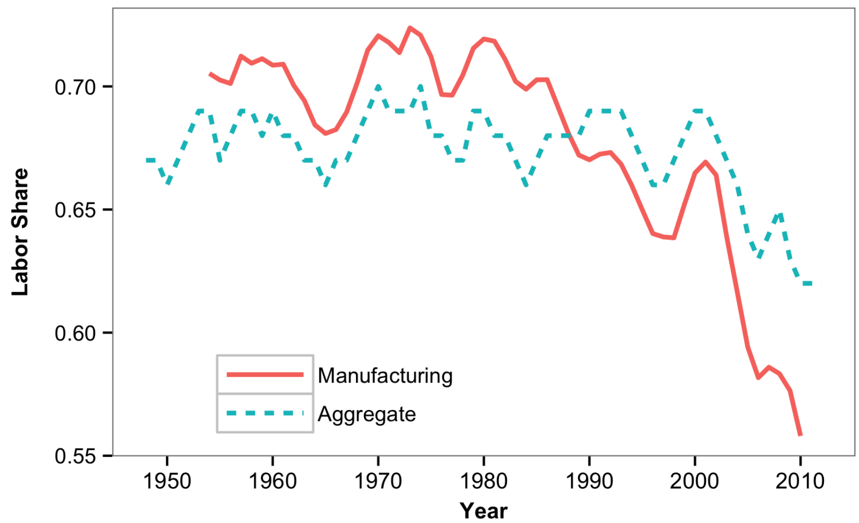
Figure 5 depicts the labor share of income for the manufacturing sector and aggregate economy in the post-war period.[29] The labor share for manufacturing has fallen since 1970, from about 0.73 to 0.55 in 2011. The largest decline has been since 2000; the labor share fell from roughly 0.65 to 0.55 in one decade. The labor share has fallen for the overall economy as well, though not by as much, falling from 0.70 in 1970 to 0.62 in 2011.

Labor's share of income could change in response to changes in technology or factor prices. The aggregate elasticity determines the impact of changes in factor prices; we ascribe the residual to the bias of technical change. We now measure the bias of technical change, and examine how it has varied over time.[30]

[29]The labor share for manufacturing comes from the BLS Multifactor Productivity Series, while the aggregate labor share comes from Fernald (2012); both are based on data from the National Accounts. As Gomme and Rupert (2004), Krueger (1999), and Elsby et al. (forthcoming) point out, the major issue with calculating the labor share is whether proprietors' income accrues to labor or capital. Both series assume that the share of labor for proprietors is the same as for corporations. For the manufacturing sector, proprietors' income represented 1.4 percent of income on average since 1970.

[30]See Appendix D.4 for the details of this calculation. Total income can be decomposed into payments to labor,

Figure 5 Labor Share over Time



Note: The solid red line is the labor share for manufacturing based on data from the BLS Multifactor Productivity series. The dashed, blue line is the labor share for the overall economy from Fernald (2012).

Formally, let $s^{v,L}$ denote labor's share of value added. Then the change in the labor share can be written as

$$ds^{v,L} = \frac{\partial s^{v,L}}{\partial \ln w/r} d \ln w/r + \left(ds^{v,L} - \frac{\partial v^L}{\partial \ln w/r} d \ln w/r \right)$$

The first measures the contribution of changes in factor prices. The second measures a residual that, through the lens of an aggregate production function, would be viewed as the bias of technical change.

To execute this decomposition, we need measures of factor prices. We base our rental prices on NIPA deflators for equipment and structures, and wages on average compensation per hour worked adjusted for changes in worker quality over time.[31] Our measures of value added and input

payments to fixed capital, and a residual that we label "profit". Any return to intangible capital or entrepreneurship would be included in this category. Consistent with Section 2, we assume changes in factor prices do not alter the "profit" share of gross output. We pursue an alternative approach in Appendix D.4.2 which assumes that factor prices (the wage and cost of capital) do not impact the split of income between fixed capital and intangible capital. In that case, the contribution from factor prices rises slightly by about 0.1 percentage points annually.

[31] For rental prices, we develop a Tornqvist index for the rental price to account for changing shares of two digit industries and different types of capital over time. For wages, we use BLS data on total compensation and hours for each industry, correcting for labor quality using indices from Jorgenson et al. (2013). Jorgenson et al. (2013) measures labor quality as the deviation of total hours from a Tornqvist index of hours across many different cells that represent workers with different amounts of human capital, as in Jorgenson et al. (2005).

expenditures are based on NIPA data. Our estimates of expenditures on fixed capital combine NIPA data on equipment and structures capital combined with our rental prices for the fixed capital contribution.[32] Finally, we allow the aggregate elasticity to vary across time, linearly interpolating between the Census years in which we estimated the elasticity.

Table III displays the annualized change in the labor share and each of its components before and after the year 2000. This type of decomposition cannot provide a full explanation of the decline in the labor share. Nevertheless, any explanation should be consistent with the patterns that we depict.

The contribution of changes in factor prices has been approximately constant and small over the past 40 years. The factor price contribution *raises* the labor share by about 0.05 percentage points per year. This casts doubt on explanations that work through changes in prices, such as an acceleration in investment specific technical change or the decline in labor supplied by prime-aged males.[33]

Instead, there has been an acceleration in the bias of technical change over the past 40 years; the labor share decreases about half a percentage point faster in the period since 2000 accounted for by the bias of technical change. These may stem from different sources, including automation, IT investment, the decline of unions, or offshoring.

From 1970-2011, annual growth in real wages in the manufacturing sector was 1.9 percentage

[32]The labor share from NIPA is based on firm level data, rather than production level data, and so would include non-manufacturing establishments of a manufacturing firm, such as a firm's headquarters. In 1987, the total wages and salaries from the production data (from the NBER-CES Productivity database) was 88.2 percent of the wages and salaries from NIPA. While our aggregate elasticities are estimated using production data, we apply these elasticities to labor shares from NIPA. This is our preferred estimate. See Appendix D.4.3 for details and alternative measures.

[33]Our rental prices are based upon official NIPA deflators for equipment and structures capital. However, Gordon (1990) has argued that the NIPA deflators underestimate the actual fall in equipment prices over time. In Appendix D.4.4 we use an alternative rental price series for equipment capital that Cummins and Violante (2002) developed by extending the work of Gordon (1990). Their series extends to 1999, so we compare our baseline to these rental prices during the 1970-1999 period. Using the Cummins and Violante (2002) equipment prices implies that the wage to rental price ratio has increased by 3.8 percent per year, instead of 2.0 percent per year with the NIPA deflators. This change increases the contribution of factor prices to the labor share between 1970 and 1999 by 0.05 percentage points per year (and also increases the magnitude of the contribution of bias by the same amount).

Table III Contributions to Labor Share Change

Period	Labor Share Change	Contributions from: Factor Prices	Bias
1970-1999	-0.25%	0.07%	-0.32%
2000-2010	-0.79%	0.05%	-0.84%

Note: All changes are annualized. The factor price and bias contributions are as defined in the text.

points lower than 1954-1969. How much of the decline in the labor share can be explained by the slowdown in wage growth? If wage growth had continued to grow at the same robust pace, how much higher would the labor share be?

We can use our estimated aggregate elasticity to assess this counterfactual.[34] We find that if wages grew faster, the manufacturing labor share in 2010 would be 0.58, compared to its actual value of 0.55. Thus the slowdown in wages can explain about 15 percent of the total decline.

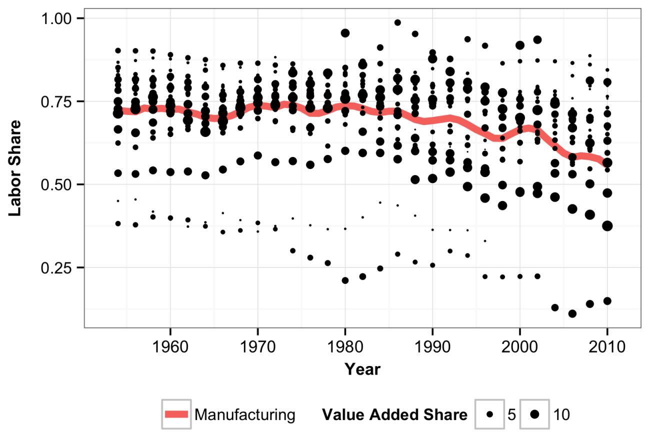
5.1 Industry Decomposition

What drives the bias of technical change? One possibility is that industries have become more capital intensive. Alternatively, the composition of industries within manufacturing could have shifted. Changes in trade barriers or the cost of outsourcing, for example, could change the aggregate capital intensity by shifting industry composition.

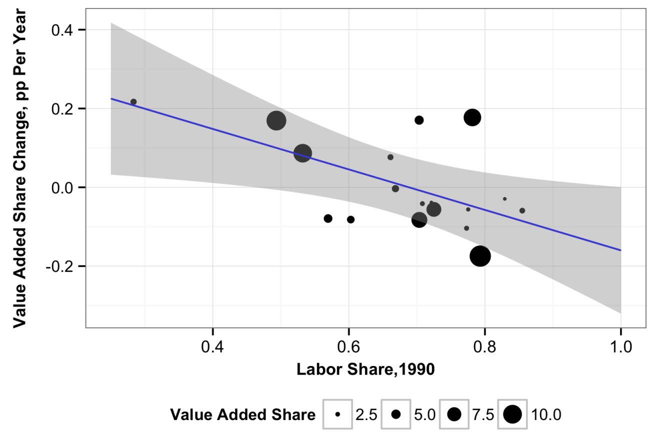
Figure 6 shows that industries that were more labor intensive tended to shrink relative to capital intensive industries. Thus shifts in composition likely played a role in the labor share's decline.

We can go further and decompose the sources of aggregate bias between biased technical change within industries and compositional change across industries. Formally, let $s_n^{v,L}$ denote labor's share of industry n's value added, and let v_n denote industry n's share of total value added. Since $s^{v,L} = \sum_n v_n s_n^{v,L}$, the change in labor's share of income can be decomposed to into within- and

[34] Note that this counterfactual involves extrapolating our local estimate of the aggregate elasticity of substitution.



(a) Industry Labor Shares



(b) Change in Value Added Shares

Figure 6 Industry Labor Shares and Value Added

Note: Panel (a) shows the evolution of industry labor shares, where the area of each point equals that industry's share of value added. Panel (b) shows each industries changes in value added shares and its average labor share of income in between 1990 and 2010.

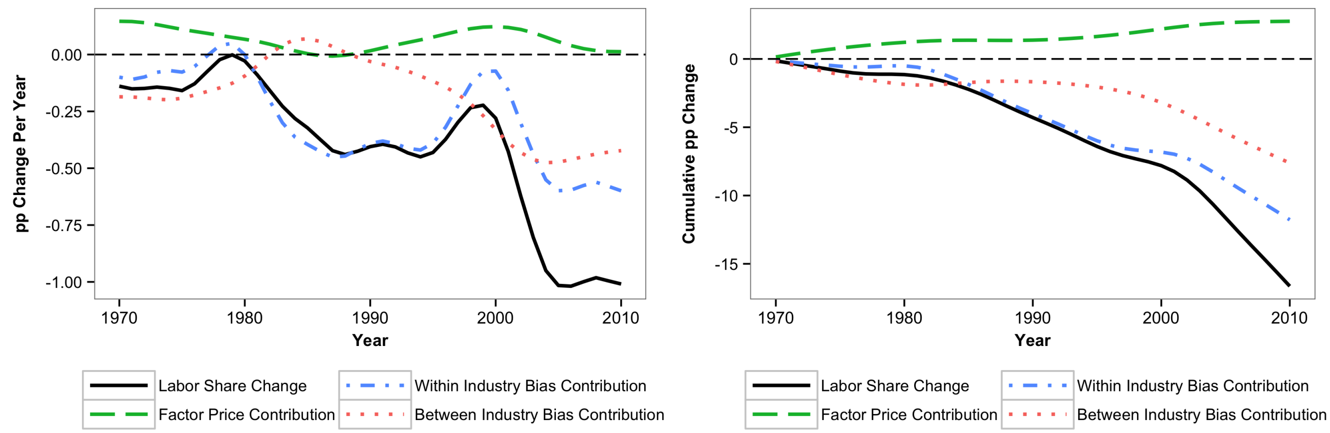
between-industry components

$$ds^{v,L} = \sum_n v_n ds_n^{v,L} + \sum_n \left(s_n^{v,L} - s^{v,L}\right) dv_n$$

The change in the overall labor can thus be decomposed into three terms:

$$ds^{v,L} = \frac{\partial s^{v,L}}{\partial \ln w/r} d\ln w/r + \sum_n v_n \left(ds_n^{v,L} - \frac{\partial s_n^{v,L}}{\partial \ln w/r} d\ln w/r \right) + \sum_n \left(s_n^{v,L} - s^{v,L}\right) \left(dv_n - \frac{\partial v_n}{\partial \ln w/r} d\ln w/r \right) \tag{12}$$

The first term measures the contribution of factor price changes. The second term measures within-industry contributions to the bias of technical change. It is the weighted average of the changes in industries' labor shares dues to technical change. The third term measures the between-industry contribution to the bias of technical change. It is the covariance between industry growth due to technical change and its labor share of income. When the covariance is positive, labor intensive industries are growing relative to capital intensive industries, which raises labor's share of value



(a) Annual Change

(b) Cumulative Change

Figure 7 Change in Labor Share Components over Time

Note: All changes are in percentage points. Each contribution to the labor share change has been smoothed using a Hodrick-Prescott filter. The contributions from factor prices, within-industry bias, and between-industry bias are as defined in the text.

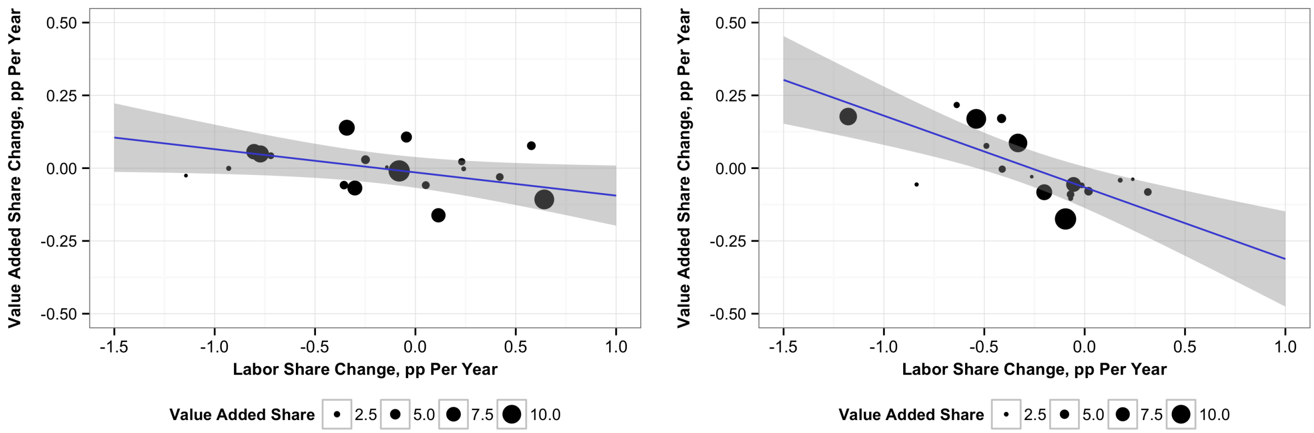
added.

Figure 7 shows the contribution of the within industry bias and between industry bias as in equation (12). Figure 7a shows the annualized rate of change while Figure 7b shows the cumulative change.[35]

This decomposition points to the importance of both technical change within industries and compositional changes in the decline of the labor share of manufacturing. The within industry bias contribution has been large since the 1980s; overall, it is responsible for about 12 percentage points of the decline in the labor share. The between industry bias was low for much of the sample period, but rose in the 2000s; overall, it contributes about 37 percent of the overall bias of technical change.

Finally, we can look at how changes within industries are related to changes across industries. Figure 8 plots each industry's annual change in value added against the annual change in its labor share. There was a clear shift in this relationship between the first and second halves of the sample

[35] We smooth the level of the labor share and each contribution by applying a Hodrick-Prescott filter; the change in the labor share is then the sum of all the components. We use a smoothing parameter of 60.



(a) 1970-1990

(b) 1990-2010

Figure 8 Changes in Value Added Shares and Changes in Labor Shares

Note: All changes are annualized and are in percentage points.

period. From 1970-1990, the two were uncorrelated, whereas after 1990 the industries that grew more were those with the largest decline in labor share. While this is consistent with automation, IT investment, or outsourcing of labor intensive tasks, it is harder to reconcile with a story in which the labor-intensive segment of an industry moves abroad.

5.2 The Aggregate Time Series Approach

We now compare our methodology to the approach that jointly estimates the aggregate capital-labor elasticity of substitution and bias of technical change using aggregate time series data. This approach uses the following econometric model:

$$\frac{s^{v,L}}{1-s^{v,L}} = \beta_0 + (\sigma^{agg} - 1)\ln\frac{r}{w} + \ln\phi + \epsilon \tag{13}$$

where $d\ln\phi$ is the bias of technical change and ϵ is interpreted as measurement error that is orthogonal to $\ln\frac{r}{w}$. It is well known that estimates depend critically on what assumptions are placed on the bias of technical change. Under an assumption of Hicks neutral technical change

($d\ln\phi = 0$), the aggregate elasticity is precisely estimated at 1.91. The elasticity is considerably above one because the labor share fell and wages rose relative to capital prices during the sample period.

Once we allow for biased technical change, however, estimates of both the bias and aggregate elasticity become imprecise, as shown in Figure 9. The first way we introduce biased technical change is through a constant rate of biased technical change ($d\ln\phi$ is constant). This constant rate of bias becomes a time trend in the aggregate regression. The elasticity is then identified by movements in relative factor prices around the trend; short run movements in factor prices are assumed to be uncorrelated with movements in technology. Given a constant bias, the estimate of the aggregate elasticity using least squares regressions is 0.56; the 95 percent confidence interval ranges from 0.05 to 1.07.

Our evidence for a rising rate of biased technical change over time motivates the use of a more flexible specification for the bias. We use a Box-Cox transformation of the time trend, as in Klump et al. (2007), which allows the bias to vary monotonically over time.[36] With the Box-Cox specification, the aggregate elasticity is 0.69, close to our baseline estimates. Again, the range of the confidence interval is large.

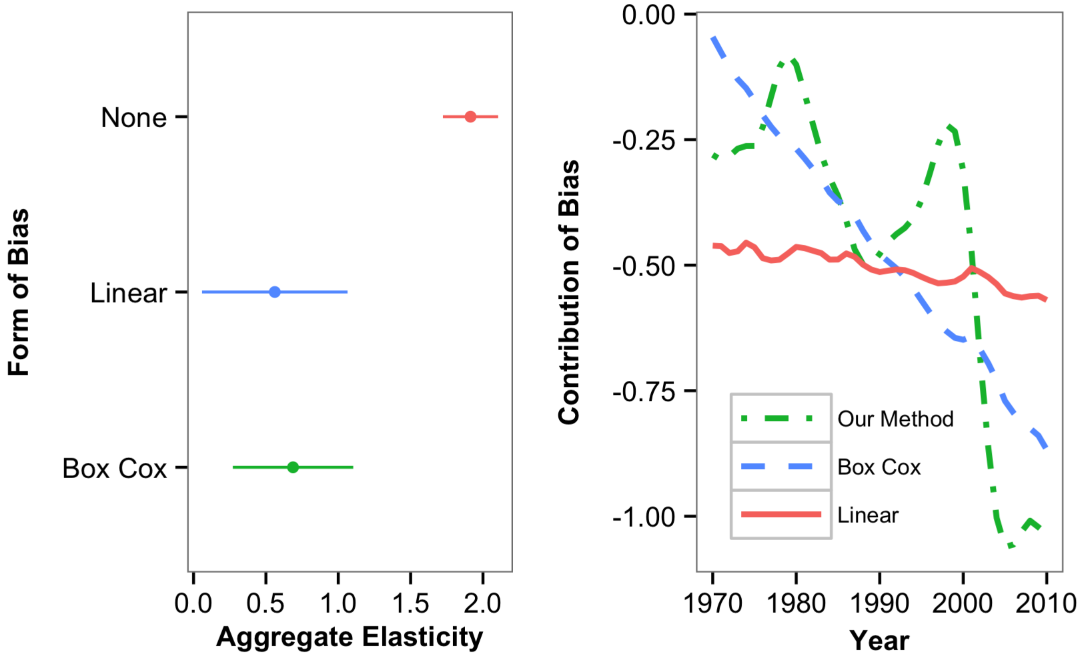
Each methodology provides a measure of the contribution of the bias of technical change to the decline in the labor share, depicted in Figure 9.[37] Assuming a constant rate of biased technical change, the average contribution of bias is about -0.5 percentage points per year and is larger than our average contribution. More importantly, this average papers over the large changes in the contribution of bias over time. The Box-Cox specification implies that the contribution of bias to the labor share has accelerated over time, but does not display the sharp drop at 2000 that the

[36]The Box-Cox transformation implies that $d\ln\phi = \gamma t^{\lambda}$; λ allows the rate of biased technical change to vary over time.

[37]For the aggregate time series method, the contribution of bias is $s^{v,L}(1-s^{v,L})d\ln\phi$; thus, the contribution to the labor share can vary over time even if the bias is constant. The contribution of bias from our method is the sum of the contributions from capital-labor bias and profit bias.

bias estimates from our method have.

Figure 9 Elasticity and Bias Estimates from Aggregate Data



Note: The left plot displays the point estimate and 95 percent confidence interval for the aggregate elasticity of substitution from regressions based on equation (13). Specifications differ in assumptions on the bias of technical change. Technical change is respectively assumed to have no trend, follow a linear time trend, or follow a Box-Cox transformation of the time trend. The right plot displays the contribution to the labor share from the bias of technical change, from either aggregate regressions with a linear or Box-Cox specification of the time trend or from our method that estimates the aggregate elasticity from the micro data. The contribution of bias from our method is the sum of the contributions of capital-labor bias and profit bias.

6 Cross–Country Elasticities

How does the aggregate capital-labor elasticity vary across countries? Production technologies may differ across countries for a number of reasons. Researchers have generally found greater variation in capital intensity and productivity in less developed countries (Hsieh and Klenow (2009), Bartlesman et al. (2009)). In addition, if lower wages in less developed countries reduce adoption of new capital intensive technologies, as in Acemoglu and Zilibotti (2001), developing countries would operate a mix of old and new technologies. In our framework, greater microeconomic heterogeneity implies a higher aggregate capital-labor elasticity of substitution.

6.1 Estimates

We examine the aggregate elasticity for three less developed countries – Chile, Colombia, and India – using micro datasets on production for each country. Appendix C.7 contains the details of each dataset; on average, there are roughly 5,000 plants per year for Chile from 1986–1996, 7,000 for Colombia from 1981–1991, and 30,000 for India from 2000–2003. In computing the aggregate elasticity, we allow the composition of plants to change across countries but fix production and demand elasticities at their US 1987 values for matching industries. Figure 10 depicts these estimates; we find an aggregate manufacturing elasticity of 0.81 in Chile, 0.84 in Colombia, and 1.15 in India, all of which are higher than the US 1987 value of 0.70.

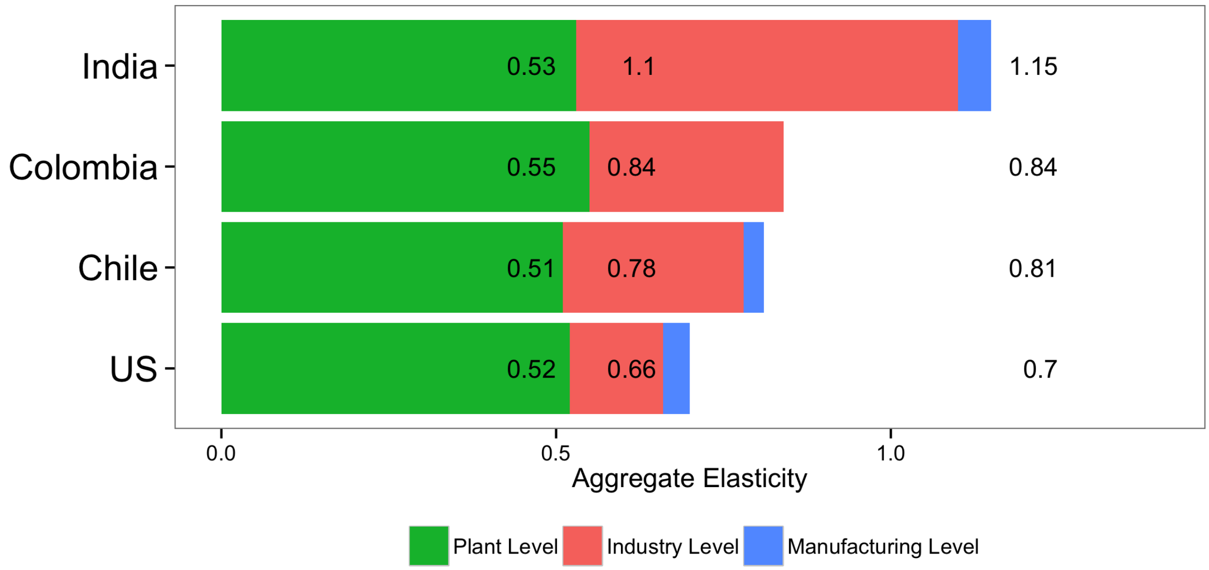
Why are aggregate elasticities higher in the less developed countries? Figure 10 depicts the average plant level and industry level elasticities; most of the difference across countries is due to a larger industry level elasticities. For India and Chile, greater heterogeneity in capital intensities accounts for roughly 70 percent of the difference in aggregate elasticities.[38] The higher elasticity for Colombia is due to a combination of multiple factors, including a different industrial composition, more heterogeneity, and lower materials shares.

6.2 Policy Implications

These cross-country differences in elasticities can imply large differences in outcomes. We examine two potential policy changes that would affect output per worker through a change in the capital rental rate. The first policy change lowers foreign interest rates to the US real interest rate; the second sets corporate taxes to zero. Our approach allows us to examine the effect of each policy

[38] Across industries, the average heterogeneity index in Chile is 0.15, 50 percent higher than the US 1987 value, and 0.30 in India, three times the US value.

Figure 10 Manufacturing Elasticity of Substitution Across Countries



Note: This figure records the average plant level, industry level, and manufacturing level elasticities of substitution. The elasticities for each country is an average over the relevant time period, which is 1986-1996 for Chile, 1981-1991 for Colombia, 2000-2003 for India, and 1987 for the US.

change without assuming that each country shares the same aggregate technology.

The capital rental rate is composed of the real interest rate R, depreciation rate δ, and effective corporate tax rate τ:

$$r = \frac{R+\delta}{1-\tau}$$

Table IV contains these rental rates; see Appendix C.7 for the details of the construction of each variable. The real interest rate is higher than the US for all three developing countries, with an interest rate differential of 1.8 percentage points for Chile, 1.9 percentage points for India, and 5.3 percentage points for Colombia. Higher real interest rates imply that all three countries have higher capital rental rates than the US; the Chilean capital rental rate is 9 percent higher than the US rate, the Indian rate 17 percent higher, and the Colombian rate 51 percent higher. Capital tax rates are roughly similar for all countries.

Table IV Cross Country Differences in Real Interest Rates, Effective Corporate Taxes, and Capital Rental Rates

	Chile	Colombia	India	US
Real Interest Rate	5.9%	9.5%	6.0%	4.2%
Effective Corporate Tax Rate	15.1%	24.3%	20.3%	18.2%
Capital Rental Rate	18.1%	25.0%	19.4%	16.6%

Note: This table records the average real interest rate from 1992–2011 using the private sector lending rate from the IMF International Financial Statistics adjusted for inflation using the change in the GDP deflator, the one year effective corporate tax rate from Djankov et al. (2010), and the capital rental rate given the average real interest rate, corporate tax rate, and a depreciation rate of 9.46 percent.

Table V displays the change in output per worker from both policy changes, as well as the change in output per worker if all countries shared the US aggregate elasticity. The elasticity of output per worker to relative factor prices is proportional to the aggregate elasticity:

$$\frac{d \ln Y/L}{d \ln r/w} = \alpha \sigma^{agg}$$

where α is the manufacturing capital share.[39]

If all interest rates fell to the US interest rate, output per worker would increase by 5.3 percent in Chile, 10.5 percent in Colombia, and 8.8 percent in India. These effects are exacerbated because the aggregate elasticity of substitution is larger in these countries than in the US. For example, the impact of the interest rate differential in India is more than fifty percent larger than it would be if India had the same elasticity of substitution as the US.

The story is similar with corporate tax rates. Reducing the corporate tax rate to zero would raise output per worker by 3.4 percent for the US, 7.0 percent for Chile, 8.9 percent for Colombia, and 15.4 percent for India; with the US aggregate elasticity, this change becomes 6.1 percent for Chile, 7.4 percent for Colombia, and 9.4 percent for India.[40] Differences in aggregate technology

[39] Two notes: First, because the elasticity that we estimate is local, counterfactual predictions for non-local changes in rental rates involves extrapolating from our local elasticity. Second, we hold the wage fixed in these experiments; the change in wage induced from changes in rental rates depends upon labor supply as well as demand. Our estimates are a lower bound on the full effects of removing interest rate differentials, while the wage change for the capital tax change depends upon whether the revenue loss is compensated for by other tax changes.

[40] We cannot interpret the change in output per worker from these policy changes as a gain in welfare because

across countries have large effects on the outcome of these policy changes.

Table V Change in Output per Worker from Policy Changes Affecting the Capital Rental Rate

Policy Change	Chile	Colombia	India	US
Equalize to US Interest Rate	5.3%	10.5%	8.8%	0%
Equalize to US Interest Rate, using US Elasticities	4.6%	8.7%	5.4%	0%
Zero Corporate Tax	7.0%	8.9%	15.4%	3.4%
Zero Corporate Tax, using US Elasticities	6.1%	7.4%	9.4%	3.4%

Note: This table records the change in output per worker from two policy experiments- equalizing all interest rates to the US interest rate and setting the corporate tax rate to zero. For each experiment, we examine the change in output per worker using the country elasticity and using the US 1987 elasticity.

7 Conclusion

This paper has developed a new approach to estimate the aggregate elasticity of substitution between capital and labor by building up from micro structural parameters and the cross-sectional distribution of plant-level expenditures. Our approach has several advantages. We can estimate the aggregate elasticity at each point in time, which means we can separately examine the bias of technical change. In addition, with cross-sectional micro data, we can examine the robustness of our model and identification strategy.

We then applied our methodology to data on plants in the US manufacturing sector. While we estimated an average plant-level elasticity of roughly 0.5, our estimates indicated that the aggregate elasticity of substitution between capital and labor has been close to 0.7 from 1972 to 2007. Thus roughly 65 percent of aggregate factor substitution happened within plants.

We then measured the contributions of factor prices and two components of the bias of technology to the fall of the labor share in manufacturing since 1970. To do this, we used our estimates of aggregate elasticities and the history of factor prices. We found that the bias of technical change within industries has steadily decreased the labor share, accounting for most of the shift. There

that comparison contrasts two different steady states. However, Chamley (1981) found in a full general equilibrium model that the welfare cost of capital taxation is proportional to the elasticity of substitution.

has been a sharp drop in the labor share since 2000 that coincides with an increase in the between-industry bias. We found that changes in factor prices do not account for the decline in the labor share.

Our work indicates that explaining the decline in the labor share requires understanding the determinants of technical change, broadly defined, including automation, ICT investment, offshoring, or shifts in composition driven by changing tastes or trade. Our decomposition suggests that there may be more than one force at work. While we have applied our methodology to the change in the US labor share, we believe it can be useful in understanding the evolution of skill premia and inequality.

Finally, our approach allowed us to estimate how the aggregate elasticity varies across time and countries, and to understand the underlying reasons for such variation. In particular, a greater heterogeneity implies a higher aggregate elasticity of substitution. Elasticities in all three of the developing countries we examined were higher than the US, with an average manufacturing sector elasticity of 0.81 for Chile, 0.84 for Colombia, and 1.15 for India. Our estimates can help illuminate how and why policies have differential effects across countries.

References

Acemoglu, Daron, "Directed Technical Change," *Review of Economic Studies*, 2002, *69* (4), 781–809.

_ , "Labor- And Capital-Augmenting Technical Change," *Journal of the European Economic Association*, 2003, *1* (1), 1–37.

_ , "When Does Labor Scarcity Encourage Innovation?," *Journal of Political Economy*, 2010, *118* (6), 1037–1078.

_ **and Fabrizio Zilibotti**, "Productivity Differences," *Quarterly Journal of Economics*, 2001, *116* (2), 563–606.

Antras, Pol, "Is the US Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," *Contributions to Macroeconomics*, 2004, *4* (1).

Atalay, Enghin, "How Important Are Sectoral Shocks?," 2014. Mimeo.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, forthcoming.

_ **, Frank Levy, and Richard J. Murmane**, "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 2003, *118*, 1279–1333.

Banerjee, Abhijit and Esther Duflo, "Growth Theory Through The Lens Of Development Economics," *Handbook of Economic Growth*, 2005, *1*, 473–552.

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

Bartlesman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocation and Selection," Technical Report, NBER Working Paper No. 1549 2009.

Berndt, Ernst R., "Reconciling Alternative Estimates of the Elasticity of Substitution," *Review of Economics and Statistics*, 1976, *58* (1), 59–68.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, "Accounting for Changes in Between-Group Inequality," 2014. Mimeo.

Caballero, Ricardo, Eduardo Engel, and John Haltiwanger, "Plant Level Adjustment and Aggregate Investment Dynamics," *Brookings Papers on Economic Activity*, 1995, *2*, 1–54.

Chamley, Christophe, "The Welfare Cost of Capital Income Taxation in a Growing Economy," *Journal of Political Economy*, 1981, *89* (3), 468–96.

Chirinko, Robert S., "Sigma: The Long and Short of It," *Journal of Macroeconomics*, 2008, *30*, 671–686.

_ **, Steven M. Fazzari, and Andrew P. Meyer**, "A New Approach to Estimating Production Function Parameters: The Elusive Capital-Labor Substitution Elasticity," *Journal of Business and Economic Statistics*, 2011, *29* (4), 587–594.

Cummins, Jason and Giovanni Violante, "Investment-Specific Technical Change in the US (1947-2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics*, 2002, *5* (2), 243–284.

Diamond, Peter, Daniel McFadden, and Miguel Rodriguez, "Measurement of the Elasticity of Factor Substitution and Bias of Technical Change," in Melvyn Fuss and Daniel McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications*, 1978, chapter 5.

Djankov, Simeon, Tim Ganser, Caralee McLiesh, Rita Ramalho, and Andrei Schliefer, "The Effect of Corporate Taxes on Investment and Entrepreneurship," *American Economic Journal: Macroeconomics*, 2010, *2* (3), 31–64.

Dornbusch, Rudiger, Stanley Fischer, and Paul Samuelson, "Heckscher-Ohlin Trade Theory with a Continuum of Goods," *Quarterly Journal of Economics*, 1980, *95* (2), 203–224.

Elsby, Michael W. L., Bart Hobijn, and Aysegul Sahin, "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity*, forthcoming.

Fernald, John, "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," September 2012. Mimeo.

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

Gomme, Paul and Peter Rupert, "Measuring Labor's Share of Income," *American Economic Review*, 2004.

Gordon, Robert J., *The Measurement of Durable Goods Prices*, University of Chicago Press, 1990.

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

Hall, Robert E. and Dale W. Jorgenson, "Tax Policy and Investment Behavior," *American Economic Review*, 1967, *57* (3), 391–414.

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.

Hicks, John R., *The Theory of Wages*, MacMillan and Co, 1932.

Houthakker, Hendrik, "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis," *The Review of Economic Studies*, 1955, *23* (1), pp. 27–31.

Hsieh, Chang-Tai and Peter Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Imbs, Jean and Isabelle Mejean, "Elasticity Optimism," May 2011. Mimeo.

Jones, Chad, "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, 2005, *120* (2), 517–549.

Jorgenson, Dale W., Mun S. Ho, and John D. Samuels, "A Prototype Industry-level Production Account for the United States, 1947-2010," 2013. Mimeo.

_ , _ , **and Kevin J. Stiroh**, "Growth of U.S. Industries and Investments in Information Technology and Higher Education," in "Measuring Capital in the New Economy" NBER Chapters, National Bureau of Economic Research, Inc, October 2005, pp. 403–478.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," Technical Report, National Bureau of Economic Research 2013.

Klump, Rainer, Peter McAdam, and Alpo Willman, "Factor Substitution and Factor Augmenting Technical Progress in the US," *Review of Economics and Statistics*, 2007, *89* (1), 183–192.

Krueger, Alan B., "Measuring Labor's Share," *American Economic Review*, 1999, *89* (2).

Krusell, Per, Lee Ohanian, Jose-Victor Rios-Rull, and Giovanni Violante, "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5).

Lagos, Ricardo, "A Model of TFP," *Review of Economic Studies*, October 2006, *73* (4), 983–1007.

Leon-Ledesma, Miguel A., Peter McAdam, and Alpo Willman, "Identifying the Elasticity of Substitution with Biased Technical Change," *The American Economic Review*, 2010, *100* (4), 1330–1357.

Levhari, David, "A Note on Houthakker's Aggregate Production Function in a Multifirm Industry," *Econometrica*, 1968, *36* (1), pp. 151–154.

Luttmer, Erzo G. J., "Growth with a Fixed Factor," 2012. Mimeo.

Mankiw, N. Gregory, "The Growth of Nations," *Brookings Papers on Economic Activity*, 1995, *1995* (1), 275–326.

Miyagiwa, Kaz and Chris Papageorgiou, "Endogenous Aggregate Easticity of Substitution," *Journal of Economic Dynamics and Control*, 2007, *31* (9), 2899–2919.

Piketty, Thomas, *Capital in the Twenty-First Century*, Belknap Press, 2014.

Raval, Devesh, "Non Neutral Technological Differences," 2013. Mimeo.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Sato, Kazuo, "A Two-Level Constant-Elasticity-of-Substitution Production Function," *The Review of Economic Studies*, 1967, *34* (2), 201–218.

Tybout, James R. and Mark Roberts, *Industrial Evolution in Developing Countries: Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure*, Oxford University Press, 1996.

Appendix

A Proofs of Propositions

Appendix A describes the aggregate elasticity of substitution between capital and labor when each plant's production function exhibits constant returns to scale. Web Appendix B derives expressions for the aggregate elasticity under alternative assumptions. Appendix B.1 characterizes local elasticities of substitution and Appendix B.2 derives preliminary results under the assumption that plants' production functions are homothetic. The assumption of constant returns to scale is relaxed in Appendix B.3. Appendix B.4 introduces misallocation frictions. Appendix B.5 generalizes the demand system to allow for arbitrary elasticities of demand and imperfect pass-through. Appendix B.6 relaxes the assumption that production functions are homothetic.

Throughout, we use the following notation for relative factor prices:

$$\begin{aligned} \omega &\equiv \frac{w}{r} \\ \mathfrak{q} &\equiv \frac{q}{r} \end{aligned}$$

In addition, we define $p_{ni} \equiv P_{ni}/r$ and $p_n \equiv P_n/r$ to be plant i's and industry n's prices respectively normalized by the rental rate. It will also be useful to define plant i's cost function (normalized by r) to be

$$z_{ni}(Y_{ni}, \omega, \mathfrak{q}) = \min_{K_{ni}, L_{ni}, M_{ni}} K_{ni} + \omega L_{ni} + \mathfrak{q} M_{ni} \qquad \text{subject to} \qquad F_{ni}(K_{ni}, L_{ni}, M_{ni}) \geq Y_{ni}$$

and industry n's cost as $z_n = \sum_{i \in I_n} z_{ni}$. Two results will be used repeatedly. First, Shephard's lemma implies that for each i:

$$(1 - s_{ni}^M)(1 - \alpha_{ni}) = \frac{\partial \ln z_{ni}}{\partial \ln \omega} \tag{14}$$

$$s_{ni}^M = \frac{\partial \ln z_{ni}}{\partial \ln \mathfrak{q}} \tag{15}$$

Second, since $\alpha_n = \sum_{i \in I_n} \alpha_{ni} \theta_{ni}$, then for any quantity κ_n,

$$\sum_{i \in I_n} (\alpha_{ni} - \alpha_n) \kappa_n \theta_{ni} = 0 \tag{16}$$

When production functions take the functional form in defined in Assumption 1, it is relatively straightforward to derive the industry-level elasticity of substitution. However, when we aggregate across industries, we have no justification for using any particular functional form for the industry-level production function, or how capital, labor, and materials are nested in the industry-level production function. We will therefore derive a formula for the industry-level elasticity of substitution in a way that nests nicely: the same formula used to go from plant-level to industry-level can be used to go from industry level to aggregate.

Consequently here we relax the functional form assumption on plants' production functions. While we maintain that each plant's production function has constant returns to scale, we impose no further structure.

Assumption 1′ *Plant i in industry n uses the constant returns to scale production function $F_{ni}(K_{ni}, L_{ni}, M_{ni})$.*

In particular, we do not assume substitution elasticities are constant, that inputs are separable, or that technological differences across plants are factor augmenting. Rather, we relate the industry-level elasticity of substitution to local elasticities of individual plants.

The purpose of this is twofold. First, it provides guidance on the robustness of the formulas in Proposition 1. Second, we will later use the formulas here to aggregate across industries and derive an expression for the elasticity of substitution between capital and labor for the manufacturing sector as a whole.

For plant i, we define the local elasticities σ_{ni} and ζ_{ni} to satisfy

$$\sigma_{ni} - 1 \quad = \quad \frac{d \ln K_{ni}/\omega L_{ni}}{d \ln \omega} \tag{17}$$

$$\zeta_{ni} - 1 \quad = \quad \frac{1}{\alpha^M - \alpha_{ni}} \frac{d \ln \frac{\mathfrak{q} M_{ni}}{K_{ni} + \omega L_{ni}}}{d \ln \omega} \tag{18}$$

If i's production function takes a nested CES form as in Assumption 1, σ_{ni} and ζ_{ni} would equal σ_n and ζ_n respectively.[41] Here, however, these elasticities are not parameters of a production function. Instead, they are defined locally in terms of derivatives of F_{ni} evaluated at i's cost-minimizing input bundle. Exact expressions for σ_{ni} and ζ_{ni} are given in Appendix B.1 of the web appendix. Proposition 1′ then expresses the industry elasticity of substitution σ_n^N. The resulting expression is identical to Proposition 1 except the plant elasticities of substitution are replaced with weighted averages of the plants' local elasticities, $\bar{\sigma}_n$ and $\bar{\zeta}_n$.

Proposition 1′ *Under Assumption 1′, the industry elasticity of substitution $\sigma_n^N = \frac{d \ln K_n/L_n}{d \ln \omega}$ is:*

$$\sigma_n^N \quad = \quad (1 - \chi_n)\bar{\sigma}_n + \chi_n \left[(1 - \bar{s}_n^M)\varepsilon_n + \bar{s}_n^M \bar{\zeta}_n \right] \tag{19}$$

where $\chi_n \equiv \sum_{i \in I_n} \frac{(\alpha_{ni} - \alpha_n)^2 \theta_{ni}}{\alpha_n(1-\alpha_n)}$ and $\bar{s}_n^M \equiv \sum_{i \in I_n} \frac{(\alpha_{ni} - \alpha_n)(\alpha_{ni} - \alpha^M)}{\sum_{j \in I_n}(\alpha_{nj} - \alpha_n)(\alpha_{nj} - \alpha^M)} s_{ni}^M$ as in Proposition 1 and

$$\bar{\sigma}_n \quad \equiv \quad \sum_{i \in I_n} \frac{\alpha_{ni}(1 - \alpha_{ni})\theta_{ni}}{\sum_{j \in I_n} \alpha_{nj}(1 - \alpha_{nj})\theta_{nj}} \sigma_{ni}$$

$$\bar{\zeta}_n \quad \equiv \quad \sum_{i \in I_n} \frac{(\alpha_{ni} - \alpha_n)(\alpha_{ni} - \alpha^M) s_{ni}^M}{\sum_{j \in I_n} (\alpha_{nj} - \alpha_n)(\alpha_{nj} - \alpha^M) s_{nj}^M} \zeta_{ni}$$

Proof. As discussed in the text, the definitions of σ_{ni} and σ_n^N imply

$$\sigma_{ni} - 1 \quad = \quad \frac{d \ln \frac{rK_{ni}}{wL_{ni}}}{d \ln \omega} = \frac{1}{\alpha_{ni}(1 - \alpha_{ni})} \frac{d\alpha_{ni}}{d \ln \omega}, \qquad \forall i \in I_n$$

$$\sigma_n^N - 1 \quad = \quad \frac{d \ln \frac{rK_n}{wL_n}}{d \ln \omega} = \frac{1}{\alpha_n(1 - \alpha_n)} \frac{d\alpha_n}{d \ln \omega}$$

[41] The definition of σ_{ni} is straightforward but ζ_{ni} requires some elaboration. Suppose that F_{ni} takes the nested CES form of Assumption 1. Let $\lambda_{ni} \equiv \frac{1}{r} \left[(r/A_{ni})^{1-\sigma_n} + (w/B_{ni})^{1-\sigma_n} \right]^{\frac{1}{\sigma_n - 1}}$ be the marginal cost of i's capital-labor bundle. Then cost minimization implies $\frac{\mathfrak{q} M_{ni}}{K_{ni} + \omega L_{ni}} = \left(\frac{\mathfrak{q}/C_{ni}}{\lambda_{ni}} \right)^{1-\zeta_n}$. Since $\frac{d \ln \mathfrak{q}}{d \ln \omega} = 1 - \alpha^M$ and, from Shephard's Lemma, $\frac{d \ln \lambda_{ni}}{d \ln \omega} = 1 - \alpha_{ni}$, we have that $\frac{d \ln \frac{\mathfrak{q} M_{ni}}{K_{ni} + \omega L_{ni}}}{d \ln \omega} = (\zeta_n - 1)(\alpha^M - \alpha_{ni})$.

Since $\alpha_n = \sum_{i \in I_n} \alpha_{ni}\theta_{ni}$, we can differentiate to get

$$\begin{aligned}
\sigma_n^N - 1 &= \frac{1}{\alpha_n(1-\alpha_n)}\left[\sum_{i \in I_n} \frac{d\alpha_{ni}}{d\ln\omega}\theta_{ni} + \sum_{i \in I_n} \alpha_{ni}\frac{d\theta_{ni}}{d\ln\omega}\right] \\
&= \frac{\sum_{i \in I_n} \alpha_{ni}(1-\alpha_{ni})\theta_{ni}(\sigma_{ni}-1)}{\alpha_n(1-\alpha_n)} + \frac{\sum_{i \in I_n} \alpha_{ni}\frac{d\theta_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)}
\end{aligned}$$

Using the definition of $\bar{\sigma}_n$ and $\sum_{i \in I_n} \frac{d\theta_{ni}}{d\ln\omega} = 0$, we have

$$\sigma_n^N - 1 = (\bar{\sigma}_n - 1)\sum_{i \in I_n} \frac{\alpha_{ni}(1-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)} + \frac{\sum_{i \in I_n}(\alpha_{ni}-\alpha_n)\frac{d\theta_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)} \tag{20}$$

Since $\chi_n = \sum_{i \in I_n} \frac{(\alpha_{ni}-\alpha_n)^2\theta_{ni}}{\alpha_n(1-\alpha_n)}$, one can verify that $\sum_{i \in I_n} \frac{\alpha_{ni}(1-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)} = 1 - \chi_n$, which gives

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n + \frac{\sum_{i \in I_n}(\alpha_{ni}-\alpha_n)\frac{d\theta_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)} + \chi_n \tag{21}$$

We now find an expression for $\frac{d\ln\theta_{ni}}{d\ln\omega}$. θ_{ni} can be written as

$$\theta_{ni} = \frac{rK_{ni} + wL_{ni}}{\sum_{j \in I_n} rK_{nj} + wL_{nj}} = \frac{(1-s_{ni}^M)z_{ni}}{\sum_{j \in I_n}(1-s_{nj}^M)z_{nj}} = \frac{(1-s_{ni}^M)z_{ni}}{(1-s_n^M)z_n}$$

Since $d\ln\frac{1-s_{ni}^M}{s_{ni}^M} = \frac{1}{s_{ni}^M} d\ln(1-s_{ni}^M)$, the definition of ζ_n implies

$$\frac{d\ln(1-s_{ni}^M)}{d\ln\omega} = s_{ni}^M(\zeta_n - 1)(\alpha_{ni} - \alpha^M)$$

The change in i's expenditure on all inputs depends on its expenditure shares:

$$\begin{aligned}
\frac{z_{ni}}{z_n} &= \frac{rK_{ni} + wL_{ni} + qM_{ni}}{\sum_{j \in I_n} rK_{nj} + wL_{nj} + qM_{nj}} = \frac{\frac{\varepsilon_n-1}{\varepsilon_n}P_{ni}Y_{ni}}{\sum_{j \in I_n}\frac{\varepsilon_n-1}{\varepsilon_n}P_{nj}Y_{nj}} = \frac{\frac{\varepsilon_n-1}{\varepsilon_n}P_{ni}^{1-\varepsilon_n}Y_nP_n^{\varepsilon_n}}{\sum_{j \in I_n}\frac{\varepsilon_n-1}{\varepsilon_n}P_{nj}^{1-\varepsilon_n}Y_nP_n^{\varepsilon_n}} \\
&= \left(\frac{p_{ni}}{p_n}\right)^{1-\varepsilon_n}
\end{aligned}$$

The change in i's price depends on the change in its marginal cost

$$\frac{d\ln p_{ni}}{d\ln\omega} = \frac{d\ln\frac{\varepsilon_n}{\varepsilon_n-1}mc_{ni}}{d\ln\omega} = \frac{d\ln mc_{ni}}{d\ln\omega} = (1-s_{ni}^M)(1-\alpha_{ni}) + s_{ni}^M(1-\alpha^M) \tag{22}$$

Putting these pieces together, we have $\theta_{ni} = \frac{1-s_{ni}^M}{1-s_n^M}\left(\frac{p_{ni}}{p_n}\right)^{1-\varepsilon_n}$, so differentiating and using equation (16) yields

$$\begin{aligned}
\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln 1-s^M_{ni}}{d\ln\omega}+(1-\varepsilon_n)\frac{d\ln p_{ni}}{d\ln\omega}\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left\{\begin{array}{c} s^M_{ni}(\zeta_{ni}-1)(\alpha_{ni}-\alpha^M) \\ +(1-\varepsilon_n)\left[(1-s^M_{ni})(1-\alpha_{ni})+s^M_{ni}(1-\alpha^M)\right]\end{array}\right\} \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left\{s^M_{ni}(\zeta_{ni}-\varepsilon_n)(\alpha_{ni}-\alpha^M)+(1-\varepsilon_n)(1-\alpha_{ni})\right\}
\end{aligned}$$

Using the definitions of $\bar{\zeta}_n$, $\bar{s}^M_n$, and χ_n, this becomes

$$\begin{aligned}
\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[s^M_{ni}(\bar{\zeta}_n-\varepsilon_n)(\alpha_{ni}-\alpha^M)+(1-\varepsilon_n)(1-\alpha_{ni})\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\bar{s}^M_n(\bar{\zeta}_n-\varepsilon_n)(\alpha_{ni}-\alpha^M)+(1-\varepsilon_n)(1-\alpha_{ni})\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\bar{s}^M_n(\bar{\zeta}_n-\varepsilon_n)(\alpha_{ni}-\alpha_n)+(1-\varepsilon_n)(\alpha_n-\alpha_{ni})\right] \\
&= \alpha_n(1-\alpha_n)\chi_n\left[(\bar{\zeta}_n-\varepsilon_n)\bar{s}^M_n-(1-\varepsilon_n)\right]
\end{aligned}$$

Finally, we can plug this back into equation (21) to get

$$\sigma^N_n=(1-\chi_n)\bar{\sigma}_n+\chi_n\left[\bar{s}^M_n\zeta_n+(1-\bar{s}^M_n)\varepsilon_n\right]$$

■

To build up to the aggregate elasticity of substitution between capital and labor, we proceed in exactly the same way as with the industry-level elasticity. Define ζ^N_n to satisfy $(\zeta^N_n-1)(\alpha_n-\alpha^M)=\frac{d\ln\frac{1-s^M_n}{s^M_n}}{d\ln\omega}$. The claim below will give an expression for ζ_n in terms of plant level elasticities and choices.

Proposition 2′ *Under Assumption 1′, the aggregate elasticity between capital and labor is*

$$\sigma^{agg}=(1-\chi^N)\bar{\sigma}^N+\chi_N\left[\bar{s}^M\bar{\zeta}^N+(1-\bar{s}^M)\eta\right]$$

where χ^N, $\bar{\sigma}^N$, $\bar{\zeta}^N$, $\bar{s}^M$, and ζ_n^N are defined as

$$\begin{aligned}
\chi^N &= \sum_{n\in N}\frac{(\alpha_n-\alpha)^2\theta_n}{\alpha(1-\alpha)}\\
\bar{\sigma}^N &= \sum_{n\in N}\frac{\alpha_n(1-\alpha_n)\theta_n}{\sum_{n'\in N}\alpha_{n'}(1-\alpha_{n'})\theta_{n'}}\sigma_n^N\\
\bar{\zeta}^N &\equiv \frac{\sum_{n\in N}(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n s_n^M\zeta_n^N}{\sum_{n\in N}(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n s_n^M}\\
\bar{s}^M &= \frac{\sum_{n\in N}(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n s_n^M}{\sum_{n\in N}(\alpha_n-\alpha)(\alpha_n-\alpha^M)\theta_n}\\
\zeta_n^N &= \sum_{i\in I_n}\theta_{ni}\left[\zeta_{ni}\frac{s_{ni}^M(\alpha_{ni}-\alpha^M)}{s_n^M(\alpha_n-\alpha^M)}+\varepsilon_n\left\{1-\frac{s_{ni}^M(\alpha_{ni}-\alpha^M)}{s_n^M(\alpha_n-\alpha^M)}\right\}\right]
\end{aligned}$$

Proof. Note that since $\frac{z_{ni}}{z_n}=\frac{P_{ni}Y_{ni}}{P_nY_n}$, we have both $\frac{d\ln Y_n}{d\ln\omega}=\sum_{i\in I_n}\frac{z_{ni}}{z_n}\frac{d\ln Y_{ni}}{d\ln\omega}$ and $\frac{d\ln p_n}{d\ln\omega}=\sum_{i\in I_n}\frac{z_{ni}}{z_n}\frac{d\ln p_{ni}}{d\ln\omega}$. With the first, we can derive an expression for the change in cost that parallels equation (27):

$$\begin{aligned}
\frac{d\ln z_n}{d\ln\omega} &= \sum_{i\in I_n}\frac{z_{ni}}{z_n}\frac{d\ln z_{ni}}{d\ln\omega}\\
&= \sum_{i\in I_n}\frac{z_{ni}}{z_n}\left[\frac{d\ln Y_{ni}}{d\ln\omega}+(1-s_{ni}^M)(1-\alpha_{ni})+s_{ni}^M(1-\alpha^M)\right]\\
&= \frac{d\ln Y_n}{d\ln\omega}+(1-s_n^M)(1-\alpha_n)+s_n^M(1-\alpha^M)
\end{aligned}$$

With the second, we can derive an expression for the change in the price level that parallels equation (22)

$$\begin{aligned}
\frac{d\ln p_n}{d\ln\omega} &= \sum_{i\in I_n}\frac{z_{ni}}{z_n}\frac{d\ln p_{ni}}{d\ln\omega}\\
&= \sum_{i\in I_n}(1-\alpha_{ni})(1-s_{ni}^M)+s_{ni}^M(1-\alpha^M)\\
&= (1-\alpha_n)(1-s_n^M)+s_n^M(1-\alpha^M)
\end{aligned}$$

Thus following the exact logic of Proposition 1′, we have that

$$\sigma^{agg}=(1-\chi^N)\bar{\sigma}^N+\chi_N\left[\bar{s}^M\bar{\zeta}^N+(1-\bar{s}^M)\eta\right]$$

It remains only to derive the expression for ζ_n^N. Begin with $1-s_n^M=\sum_{i\in I_n}(1-s_{ni}^M)\frac{z_{ni}}{z_n}$. Differentiating each side gives:

$$\frac{d\ln(1-s_n^M)}{d\ln\omega}=\sum_{i\in I_n}\frac{(1-s_{ni}^M)z_{ni}}{(1-s_n^M)z_n}\left[\frac{d\ln(1-s_{ni}^M)}{d\ln\omega}+\frac{d\ln z_{ni}/z_n}{d\ln\omega}\right]=\sum_{i\in I_n}\theta_{ni}\left[\frac{d\ln(1-s_{ni}^M)}{d\ln\omega}+\frac{d\ln z_{ni}/z_n}{d\ln\omega}\right]$$

Using $(\alpha_{ni} - \alpha^M)(\zeta_{ni} - 1) = \frac{1}{s_{ni}^M} \frac{d \ln(1-s_{ni}^M)}{d \ln \omega}$ and $(\alpha_n - \alpha^M)(\zeta_n^N - 1) = \frac{1}{s_n} \frac{d \ln(1-s_n^M)}{d \ln \omega}$ gives

$$s_n^M(\alpha_n - \alpha^M)(\zeta_n^N - 1) = \sum_{i \in I_n} \theta_{ni} \left[s_{ni}^M(\alpha_{ni} - \alpha^M)(\zeta_{ni} - 1) + \frac{d \ln z_{ni}/z_n}{d \ln \omega} \right]$$

Finally, we have

$$\begin{aligned}
\frac{d \ln z_{ni}/z_n}{d \ln \omega} &= \frac{d \ln Y_{ni}}{d \ln \omega} + (1 - s_{ni}^M)(1 - \alpha_{ni}) + s_{ni}^M(1 - \alpha^M) - \frac{d \ln Y_n}{d \ln \omega} - (1 - s_n^M)(1 - \alpha_n) - s_n^M(1 - \alpha^M) \\
&= (-\varepsilon_n)\frac{d \ln p_{ni}/p_n}{d \ln \omega} + (1 - s_{ni}^M)(1 - \alpha_{ni}) + s_{ni}^M(1 - \alpha^M) - (1 - s_n^M)(1 - \alpha_n) - s_n^M(1 - \alpha^M) \\
&= (1 - \varepsilon_n)\left[(1 - s_{ni}^M)(1 - \alpha_{ni}) + s_{ni}^M(1 - \alpha^M) - (1 - s_n^M)(1 - \alpha_n) - s_n^M(1 - \alpha^M)\right] \\
&= (\varepsilon_n - 1)\left[s_n^M(\alpha_n - \alpha^M) - s_{ni}^M(\alpha_{ni} - \alpha^M) + (\alpha_{ni} - \alpha_n)\right]
\end{aligned}$$

Plugging this in, we have

$$s_n^M(\alpha_n - \alpha^M)(\zeta_n^N - 1) = \sum_{i \in I_n} \theta_{ni} \{ s_{ni}^M(\alpha_{ni} - \alpha^M)(\zeta_{ni} - 1) + (\varepsilon_n - 1) [s_n^M(\alpha_n - \alpha^M) - s_{ni}^M(\alpha_{ni} - \alpha^M) + (\alpha_{ni} - \alpha_n)$$

Using $\sum_{i \in I_n} \theta_{ni}(1 - \varepsilon_n)(\alpha_{ni} - \alpha_n) = 0$, dividing through by $s_n^M(\alpha_n - \alpha^M)$ and subtracting 1 from each side gives the result. ■

B Additional Analytical Results (Online Appendix)

This appendix describes the aggregate elasticity of substitution between capital and labor in a variety of environments. Appendix B.1 describes local elasticities of substitution and Appendix B.2 derives a preliminary result under the assumption each plant's production function is homothetic. The assumption of constant returns to scale is relaxed in Appendix B.3. Appendix B.4 introduces misallocation frictions. Appendix B.5 generalizes the demand system to allow for arbitrary elasticities of demand and imperfect pass-through. Appendix B.6 relaxes the assumption that production functions are homothetic.

As in Appendix A, we use the following notation for relative factor prices: $\omega \equiv \frac{w}{r}$ and $\mathfrak{q} \equiv \frac{q}{r}$. In addition, we define $p_{ni} \equiv P_{ni}/r$ and $p_n \equiv P_n/r$ to be plant i's and industry n's prices respectively normalized by the rental rate. It will also be useful to define i's cost function (normalized by r) to be

$$z_{ni}(Y_{ni}, \omega, \mathfrak{q}) = \min_{K_{ni}, L_{ni}, M_{ni}} K_{ni} + \omega L_{ni} + \mathfrak{q} M_{ni} \qquad \text{subject to} \qquad F_{ni}(K_{ni}, L_{ni}, M_{ni}) \geq Y_{ni}$$

As in Appendix A, two results will be used repeatedly. First, Shephard's lemma implies that for each i:

$$(1 - s_{ni}^M)(1 - \alpha_{ni}) = \frac{z_{ni\omega}(Y_{ni}, \omega, \mathfrak{q})\omega}{z_{ni}(Y_{ni}, \omega, \mathfrak{q})} \tag{23}$$

$$s_{ni}^M = \frac{z_{ni\mathfrak{q}}(Y_{ni}, \omega, \mathfrak{q})\mathfrak{q}}{z_{ni}(Y_{ni}, \omega, \mathfrak{q})} \tag{24}$$

Second, $\alpha_n = \sum_{i \in I_n} \alpha_{ni} \theta_{ni}$, so for any quantity κ_n,

$$\sum_{i \in I_n} (\alpha_{ni} - \alpha_n) \kappa_n \theta_{ni} = 0 \tag{25}$$

B.1 Locally-Defined Elasticities

In our baseline analysis we assumed that plant i produced using a nested CES production function of the form

$$F_{ni}(K_{ni}, L_{ni}, M_{ni}) = \left(\left[(A_{ni} K_{ni})^{\frac{\sigma-1}{\sigma}} + (B_{ni} L_{ni})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1} \frac{\zeta-1}{\zeta}} + (C_{ni} M_{ni})^{\frac{\zeta-1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}}$$

In that context, σ was i's elasticity of substitution between capital and labor and ζ was i's elasticity of substitution between materials and i's capital-labor bundle.

When i's production function does not take this parametric form, we define local elasticities of substitution. Suppose that i produces using the production function $Y_{ni} = F_{ni}(K_{ni}, L_{ni}, M_{ni})$ with corresponding cost function z_{ni}. We define σ_{ni} and ζ_{ni} to satisfy

$$\sigma_{ni} - 1 = \left. \frac{d \ln \frac{\alpha_{ni}}{1 - \alpha_{ni}}}{d \ln \omega} \right|_{Y_{ni} \text{ is constant}}$$

$$(\alpha_{ni} - \alpha^M)(\zeta_{ni} - 1) = \left. \frac{d \ln \frac{1 - s_{ni}^M}{s_{ni}^M}}{d \ln \omega} \right|_{Y_{ni} \text{ is constant}}$$

σ_{ni} and ζ_{ni} measure how i's relative factor usage changes in response to changes in relative factor prices holding i's output fixed (as one moves along an isoquant). That output remains fixed is relevant only if production functions are non-homothetic, in which case a change in a plant's scale would alter its relative factor usage. This section derives expressions for σ_{ni} and ζ_{ni} in terms of i's cost function.

Claim 1 *σ_{ni} and ζ_{ni} satisfy*

$$\begin{aligned}
(\alpha_{ni}-\alpha^M)\zeta_{ni} &= -\frac{1}{1-s^M_{ni}}\left[\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)\right]\\
\sigma_{ni} &= -\frac{1}{\alpha_{ni}}\left\{\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}(1-\alpha^M)+\frac{s^M_{ni}}{1-s^M_{ni}}\left[\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)\right]\right\}
\end{aligned}$$

Proof. Differentiating equation (24) and equation (23) with respect to ω gives

$$\begin{aligned}
\frac{d\ln s^M_{ni}}{d\ln\omega} &= \left\{\begin{array}{c}\frac{z_{ni\mathfrak{q}Y}Y_{ni}}{z_{ni\mathfrak{q}}}\frac{d\ln Y_{ni}}{d\ln\omega}+\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}\frac{d\ln\mathfrak{q}}{d\ln\omega}+\frac{d\ln\mathfrak{q}}{d\ln\omega}\\ -\frac{z_{niY}Y_{ni}}{z_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}-\frac{z_{ni\omega}\omega}{z_{ni}}-\frac{z_{ni\mathfrak{q}}\mathfrak{q}}{z_{ni}}\frac{d\ln\mathfrak{q}}{d\ln\omega}\end{array}\right\}\\
\frac{d\ln(1-s^M_{ni})}{d\ln\omega}+\frac{d\ln(1-\alpha_{ni})}{d\ln\omega} &= \left\{\begin{array}{c}\frac{z_{ni\omega Y}Y_{ni}}{z_{ni\omega}}\frac{d\ln Y_{ni}}{d\ln\omega}+\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}\frac{d\ln\mathfrak{q}}{d\ln\omega}+1\\ -\frac{z_{niY}Y_{ni}}{z_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}-\frac{z_{ni\omega}\omega}{z_{ni}}-\frac{z_{ni\mathfrak{q}}\mathfrak{q}}{z_{ni}}\frac{d\ln\mathfrak{q}}{d\ln\omega}\end{array}\right\}
\end{aligned}$$

Imposing $\frac{d\ln Y_{ni}}{d\ln\omega}=0$, $\frac{d\ln\mathfrak{q}}{d\ln\omega}=1-\alpha^M$, and Shephard's lemma, these equations can be written as

$$\begin{aligned}
\frac{d\ln s^M_{ni}}{d\ln\omega} &= \frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)+(1-\alpha^M)-(1-s^M_{ni})(1-\alpha_{ni})-s^M_{ni}(1-\alpha^M)\\
\frac{d\ln(1-s^M_{ni})}{d\ln\omega}+\frac{d\ln(1-\alpha_{ni})}{d\ln\omega} &= \frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}(1-\alpha^M)+1-(1-s^M_{ni})(1-\alpha_{ni})-s^M_{ni}(1-\alpha^M)
\end{aligned}$$

Simplifying yields

$$\begin{aligned}
\frac{d\ln s^M_{ni}}{d\ln\omega} &= \frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)+(1-s^M_{ni})(\alpha_{ni}-\alpha^M)\\
\frac{d\ln(1-s^M_{ni})}{d\ln\omega}+\frac{d\ln(1-\alpha_{ni})}{d\ln\omega} &= \frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}(1-\alpha^M)+\alpha_{ni}-s^M_{ni}(\alpha_{ni}-\alpha^M)
\end{aligned}$$

Using $\frac{d\ln 1-s^M_{ni}}{d\ln\omega}=-\frac{s^M_{ni}}{1-s^M_{ni}}\frac{d\ln s^M_{ni}}{d\ln\omega}$ and plugging the first into the second yields

$$\frac{d\ln(1-\alpha_{ni})}{d\ln\omega} = \frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}(1-\alpha^M)+\frac{s^M_{ni}}{1-s^M_{ni}}\left[\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)\right]+\alpha_{ni}$$

Finally, the definitions of the elasticities imply $\sigma_{ni}-1=-\frac{1}{\alpha_{ni}}\frac{d\ln(1-\alpha_{ni})}{d\ln\omega}$ and $(\alpha_{ni}-\alpha^M)(\zeta_{ni}-1)=-\frac{1}{1-s^M_{ni}}\frac{d\ln s^M_{ni}}{d\ln\omega}$, so that

$$\begin{aligned}
(\alpha_{ni}-\alpha^M)\zeta_{ni} &= -\frac{1}{1-s^M_{ni}}\left[\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)\right]\\
\sigma_{ni} &= -\frac{1}{\alpha_{ni}}\left\{\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}+\frac{z_{ni\omega\mathfrak{q}}\mathfrak{q}}{z_{ni\omega}}(1-\alpha^M)+\frac{s^M_{ni}}{1-s^M_{ni}}\left[\frac{z_{ni\mathfrak{q}\omega}\omega}{z_{ni\mathfrak{q}}}+\frac{z_{ni\mathfrak{q}\mathfrak{q}}\mathfrak{q}}{z_{ni\mathfrak{q}}}(1-\alpha^M)\right]\right\}
\end{aligned}$$

■

B.2 Industry Substitution and Within-Plant Substitution

We now define i's local returns to scale to be $\gamma_{ni} = \left[\frac{z_{niY}Y_{ni}}{z_{ni}}\right]^{-1}$. The next lemma characterizes the within-plant components of industry substitution.

Lemma 1 *Suppose that plant i produces using the homothetic production function F_{ni}. The industry elasticity of substitution for industry n, σ_n^N, can be written as*

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n + \chi_n \bar{s_n^M}\bar{\zeta}_n + \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)}$$

where $\bar{\zeta}_n \equiv \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)(\alpha_{ni}-\alpha^M)s_{ni}^M}{\sum_{j\in I_n}(\alpha_{nj}-\alpha_n)(\alpha_{nj}-\alpha^M)s_{nj}^M}\zeta_{ni}$ *and* $\bar{s}_n^M \equiv \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)(\alpha_{ni}-\alpha^M)}{\sum_{j\in I_n}(\alpha_{nj}-\alpha_n)(\alpha_{nj}-\alpha^M)}s_{ni}^M$

Proof. Following the steps of the proof of Proposition 1′, we have

$$\begin{aligned}
\sigma_n^N &= (1-\chi_n)\sigma_n + \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\frac{d\theta_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)} + \chi_n \qquad (26)\\
\theta_{ni} &= \frac{rK_{ni}+wL_{ni}}{\sum_{j\in I_n} rK_{nj}+wL_{nj}} = \frac{(1-s_{ni}^M)z_{ni}}{\sum_{j\in I_n}(1-s_{nj}^M)z_{nj}}\\
\frac{d\ln(1-s_{ni}^M)}{d\ln\omega} &= s_{ni}^M(\zeta_{ni}-1)(\alpha_{ni}-\alpha^M)
\end{aligned}$$

The change in i's expenditure on all inputs depends on its return to scale and its expenditure shares:

$$\begin{aligned}
\frac{d\ln z_{ni}(Y_{ni},\omega,\mathfrak{q})}{d\ln\omega} &= \frac{Y_{ni}z_{niY}}{z_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega} + \frac{z_{ni\omega}\omega}{z_{ni}} + \frac{z_{ni\mathfrak{q}}\mathfrak{q}}{z_{ni}}\frac{d\ln\mathfrak{q}}{d\ln\omega}\\
&= \frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega} + (1-s_{ni}^M)(1-\alpha_{ni}) + s_{ni}^M(1-\alpha^M)\\
&= \frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega} + (1-\alpha_{ni}) + s_{ni}^M(\alpha_{ni}-\alpha^M) \qquad (27)
\end{aligned}$$

Putting these pieces together, since $\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\sum_{j\in I_n}(1-s_{nj}^M)z_{nj}}{d\ln\omega} = 0$, we have

$$\begin{aligned}
\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln 1-s_{ni}^M}{d\ln\omega} + \frac{d\ln z_{ni}}{d\ln\omega}\right]\\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\begin{array}{c} s_{ni}^M(\zeta_{ni}-1)(\alpha_{ni}-\alpha^M) + \frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}\\ +(1-\alpha_{ni}) + s_{ni}^M(\alpha_{ni}-\alpha^M)\end{array}\right]\\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[s_{ni}^M\zeta_{ni}(\alpha_{ni}-\alpha^M) + \frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega} + (1-\alpha_{ni})\right]
\end{aligned}$$

Using the definitions of $\bar{\zeta}_n$ and $\bar{s}_n^M$, this becomes

$$\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega} = \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\bar{s}_n^M\bar{\zeta}_n(\alpha_{ni}-\alpha^M) + \frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega} + (1-\alpha_{ni})\right]$$

Using the fact that for any constant κ, $\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\kappa = 0$, we can write this as

$$\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\bar{s}_n^M\bar{\zeta}_n(\alpha_{ni}-\alpha_n)+\frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}-(\alpha_{ni}-\alpha_n)\right]$$

Finally, we can plug this back into equation (26) to get

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n+\chi_n\bar{s}_n^M\bar{\zeta}_n+\frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{1}{\gamma_{ni}}\frac{d\ln Y_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)}$$

■

B.3 Returns to Scale

This section relaxes the assumption that each plant's production function exhibits constant returns to scale.

Claim 2 *Suppose that i produces using the production function $Y_{ni}=F_{ni}(K_{ni},L_{ni},M_{ni})=G_{ni}(K_{ni},L_{ni},M_{ni})^{\gamma}$, where G_{ni} has constant returns to scale and $\gamma\leq\frac{\varepsilon_n}{\varepsilon_n-1}$. Let $x=\frac{\varepsilon_n}{\varepsilon_n+\gamma(1-\varepsilon_n)}$. Then the industry elasticity of substitution is*

$$\sigma_n=(1-\chi_n)\bar{\sigma}_n+\chi_n\left[\bar{s}_n^M\bar{\zeta}_n+(1-\bar{s}_n^M)x\right]$$

and the revenue-cost ratio is $\frac{P_{ni}Y_{ni}}{rK_{ni}+wL_{ni}+qM_{ni}}=\frac{x}{x-1}$.

Proof. i's optimal price is $p_{ni}=\frac{\varepsilon_n}{\varepsilon_n-1}z_{niY}(Y_{ni},\omega,\mathfrak{q})$, so differentiating yields

$$\frac{d\ln p_{ni}}{d\ln\omega}=\frac{z_{niYY}Y_{ni}}{z_{niY}}\frac{d\ln Y_{ni}}{d\ln\omega}+\frac{z_{niY\omega}\omega}{z_{niY}}+\frac{z_{niY\mathfrak{q}}\mathfrak{q}}{z_{niY}}\frac{d\ln\mathfrak{q}}{d\ln\omega}$$

The production function implies that $\frac{z_{niYY}Y_{ni}}{z_{niY}}=\frac{1}{\gamma}-1$, $\frac{z_{niY\omega}\omega}{z_{niY}}=(1-\alpha_{ni})(1-s_{ni}^M)$, and $\frac{z_{niY\mathfrak{q}}\mathfrak{q}}{z_{niY}}=s_{ni}^M$, so this can be written as

$$\frac{d\ln p_{ni}}{d\ln\omega}=\left(\frac{1}{\gamma}-1\right)\frac{d\ln Y_{ni}}{d\ln\omega}+(1-\alpha_{ni})(1-s_{ni}^M)+s_{ni}^M(1-\alpha^M)$$

The change in i's output is then

$$\begin{aligned}\frac{d\ln Y_{ni}}{d\ln\omega} &= -\varepsilon_n\frac{d\ln p_{ni}}{d\ln\omega}+\frac{d\ln Y_np_n^{\varepsilon_n}}{d\ln\omega}\\ &= -\varepsilon_n\left(\frac{1}{\gamma}-1\right)\frac{d\ln Y_{ni}}{d\ln\omega}-\varepsilon_n\left[(1-\alpha_{ni})(1-s_{ni}^M)+s_{ni}^M(1-\alpha^M)\right]+\left[\frac{d\ln Y_np_n^{\varepsilon_n}}{d\ln\omega}\right]\end{aligned}$$

This can be rearranged as

$$\frac{d\ln Y_{ni}}{d\ln\omega}=\gamma x\left[(\alpha_{ni}-\alpha_n)-s_{ni}^M(\alpha_{ni}-\alpha^M)\right]+\frac{x\gamma}{\varepsilon_n}\left[\frac{d\ln Y_np_n^{\varepsilon_n}}{d\ln\omega}-\varepsilon_n(1-\alpha_n)\right]$$

Using Lemma 1 and the fact that $\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln Y_n p_n^{\varepsilon_n}}{d\ln\omega}-\varepsilon_n(1-\alpha_n)\right]=0$ gives

$$\begin{aligned}
\sigma_n^N &= (1-\chi_n)\bar{\sigma}_n+\chi_n\bar{s}_n^M\bar{\zeta}_n+\frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}x\left[(\alpha_{ni}-\alpha_n)-s_{ni}^M(\alpha_{ni}-\alpha^M)\right]}{\alpha_n(1-\alpha_n)}\\
&= (1-\chi_n)\bar{\sigma}_n+\chi_n\bar{s}_n^M\bar{\zeta}_n+\frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}x\left[(\alpha_{ni}-\alpha_n)-\bar{s}_n^M(\alpha_{ni}-\alpha^M)\right]}{\alpha_n(1-\alpha_n)}
\end{aligned}$$

where the second line uses the definition of $\bar{s}_n^M$. The desired result follows using $\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\bar{s}_n^M(\alpha^M-\alpha_n)=0$ and the definition of χ_n.

Finally, since $p_{ni}=\frac{\varepsilon_n}{\varepsilon_n-1}z_{niY}$, the revenue cost ratio is

$$\frac{P_{ni}Y_{ni}}{rK_{ni}+wL_{ni}+qM_{ni}}=\frac{p_{ni}Y_{ni}}{z_{ni}}=\frac{\varepsilon_n}{\varepsilon_n-1}\frac{z_{niY}Y_{ni}}{z_{ni}}=\frac{\varepsilon_n}{\varepsilon_n-1}\frac{1}{\gamma}=\frac{x}{x-1}$$

■

B.4 Adjustment costs and Misallocation Frictions

Suppose that $\{T_{Kni},T_{Lni},T_{Yni}\}_{i\in I_n}$ represent wedges that are paid while $\{\tau_{Kni},\tau_{Lni},\tau_{Yni}\}_{i\in I_n}$ represent wedges that are unpaid. Then, for example, if w is the overall wage level then plant i pays a wage of $T_{Lni}w$, but its shadow cost of labor is $\tau_{Lni}T_{Lni}w$. Then i acts as if it maximizes

$$\tau_{Yni}T_{Yni}P_{ni}Y_{ni}-\tau_{Kni}T_{Kni}rK_{ni}-\tau_{Lni}T_{Lni}wL_{ni}$$

subject to $Y_{ni}\le F_{ni}(K_{ni},L_{ni})$ and $Y_{ni}\le Y_nP_n^{\varepsilon_n}D_{ni}P_{ni}^{-\varepsilon_n}$. We define, α_{ni}, θ_{ni}, α_n, and z_{ni} as summarizing payments to factors

$$\begin{aligned}
z_{ni} &= T_{Kni}K_{ni}+\omega T_{Lni}L_{ni}\\
\alpha_{ni} &= \frac{T_{Kni}rK_{ni}}{T_{Kni}rK_{ni}+T_{Lni}wL_{ni}}\\
\alpha_n &= \frac{\sum_{i\in I_n}T_{Kni}rK_{ni}}{\sum_{i\in I_n}T_{Kni}rK_{ni}+T_{Lni}wL_{ni}}=\sum_{i\in I_n}\alpha_{ni}\theta_{ni}\\
\theta_{ni} &= \frac{T_{Kni}rK_{ni}+T_{Lni}wL_{ni}}{\sum_{j\in I_n}T_{Knj}rK_{nj}+T_{Lnj}wL_{nj}}=\frac{z_{ni}}{\sum_{j\in I_n}z_{nj}}
\end{aligned}$$

We define the same variables with hats to summarize shadow costs:

$$\begin{aligned}
\hat{z_{ni}} &= \tau_{Kni}T_{Kni}K_{ni}+\tau_{Lni}T_{Lni}\omega L_{ni}\\
\hat{\alpha_{ni}} &= \frac{\tau_{Kni}T_{Kni}rK_{ni}}{\tau_{Kni}T_{Kni}rK_{ni}+\tau_{Lni}T_{Lni}wL_{ni}}\\
\hat{\alpha_n} &= \frac{\sum_{i\in I_n}T_{Kni}rK_{ni}}{\sum_{i\in I_n}T_{Kni}rK_{ni}+T_{Lni}wL_{ni}}
\end{aligned}$$

Claim 3 *Suppose each F_{ni} has a constant elasticity of substitution between capital and labor, σ_{ni}. The industry elasticity of substitution, which satisfies $\sigma_n^N-1=\frac{d\ln\frac{\sum_{i\in I_n}T_{Kni}K_{ni}}{\sum_{i\in I_n}T_{Lni}\omega L_{ni}}}{d\ln\omega}$, is*

$$\sigma_n^N=(1-\hat{\chi}_n)\bar{\sigma}_n+\hat{\chi}_n\varepsilon_n$$

where $\hat{\chi}_n \equiv \sum_{i\in I_n} \frac{(\alpha_{ni}-\alpha_n)(\hat{\alpha_{ni}}-\hat{\alpha_n})\theta_{ni}}{\alpha_n(1-\alpha_n)}$

Proof. To compute the industry elasticity of substitution we have

$$\frac{d\ln\alpha_{ni}}{d\ln\omega} = \frac{d\ln\frac{\frac{T_{Kni}K_{ni}}{T_{Lni}\omega L_{ni}}}{\frac{T_{Kni}K_{ni}}{T_{Lni}\omega L_{ni}}+1}}{d\ln\omega} = \frac{T_{Lni}\omega L_{ni}}{T_{Kni}K_{ni}+T_{Lni}\omega L_{ni}}\frac{d\ln\frac{T_{Kni}K_{ni}}{T_{Lni}\omega L_{ni}}}{d\ln\omega} = (1-\alpha_{ni})(\sigma_{ni}-1)$$

and similarly

$$\frac{d\ln\alpha_n}{d\ln\omega} = \frac{d\ln\left[\frac{\sum_{i\in I_n}T_{Kni}K_{ni}}{\sum_{i\in I_n}T_{Lni}\omega L_{ni}}/\left(\frac{\sum_{i\in I_n}T_{Kni}K_{ni}}{\sum_{i\in I_n}T_{Lni}\omega L_{ni}}+1\right)\right]}{d\ln\omega} = (1-\alpha_n)(\sigma_n^N-1)$$

Using $d\alpha_n = \sum_{i\in I_n}\theta_{ni}d\alpha_{ni}+\alpha_{ni}d\theta_{ni}$, we can use the same logic as the benchmark case to write

$$\sigma_n^N - 1 = \frac{1}{\alpha_n(1-\alpha_n)}\sum_{i\in I_n}\left[\alpha_{ni}(1-\alpha_{ni})(\sigma_{ni}-1)\theta_{ni}+(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega}\right] \tag{28}$$

Before computing $\frac{d\ln\theta_{ni}}{d\ln\omega}$, note that

$$\frac{z_{ni}}{\hat{z_{ni}}} = \frac{T_{Kni}K_{ni}+T_{Lni}\omega L_{ni}}{\tau_{Kni}T_{Kni}K_{ni}+\tau_{Lni}T_{Lni}\omega L_{ni}} = \tau_{Kni}^{-1}\hat{\alpha_{ni}}+\tau_{Lni}^{-1}(1-\hat{\alpha_{ni}})$$

Since $\frac{d\ln\frac{\hat{\alpha_{ni}}}{1-\hat{\alpha_{ni}}}}{d\ln\omega} = \frac{d\ln\frac{\tau_{Kni}T_{Kni}K_{ni}}{\tau_{Lni}T_{Lni}\omega L_{ni}}}{d\ln\omega} = \sigma_{ni}-1$, we can differentiate to get

$$\begin{aligned}\frac{d\ln z_{ni}}{d\ln\omega}-\frac{d\ln\hat{z_{ni}}}{d\ln\omega} &= \alpha_{ni}\frac{d\ln\hat{\alpha_{ni}}}{d\ln\omega}+(1-\alpha_{ni})\frac{d\ln 1-\hat{\alpha_{ni}}}{d\ln\omega}\\ &= \alpha_{ni}(1-\hat{\alpha_{ni}})(\sigma_{ni}-1)-(1-\alpha_{ni})\hat{\alpha_{ni}}(\sigma_{ni}-1)\\ &= (\alpha_{ni}-\hat{\alpha_{ni}})(\sigma_{ni}-1)\end{aligned}$$

Constant returns to scale and Shephard's lemma imply that

$$\frac{d\ln\hat{z_{ni}}}{d\ln\omega} = \frac{d\ln Y_{ni}}{d\ln\omega}+(1-\hat{\alpha_{ni}})$$

Thus we have

$$\begin{aligned}\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln\theta_{ni}}{d\ln\omega} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln z_{ni}}{d\ln\omega}\\ &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln\hat{z_{ni}}}{d\ln\omega}+(\alpha_{ni}-\hat{\alpha_{ni}})(\sigma_{ni}-1)\right]\\ &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln Y_{ni}}{d\ln\omega}+(1-\hat{\alpha_{ni}})+(\alpha_{ni}-\hat{\alpha_{ni}})(\sigma_{ni}-1)\right]\\ &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln Y_{ni}}{d\ln\omega}+(\hat{\alpha_n}-\hat{\alpha_{ni}})+(\alpha_{ni}-\hat{\alpha_{ni}})(\sigma_{ni}-1)\right]\end{aligned}$$

Combining this with equation (28) gives

$$\begin{aligned}
\sigma_n^N - 1 &= \frac{1}{\alpha_n(1-\alpha_n)} \sum_{i \in I_n} \left\{ \begin{array}{c} \alpha_{ni}(1-\alpha_{ni})(\sigma_{ni}-1)\theta_{ni} \\ +(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{d\ln Y_{ni}}{d\ln\omega} + (\hat{\alpha_n} - \hat{\alpha_{ni}}) + (\alpha_{ni} - \hat{\alpha_{ni}})(\sigma_{ni}-1)\right] \end{array} \right\} \\
&= \frac{\sum_{i\in I_n} \left\{\alpha_{ni}(1-\alpha_{ni}) + (\alpha_{ni}-\alpha_n)(\alpha_{ni}-\hat{\alpha_{ni}})\right\}\theta_{ni}(\sigma_{ni}-1)}{\alpha_n(1-\alpha_n)} \\
&\quad + \frac{\sum_{i\in I_n}\left\{(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}}{d\ln\omega} + (\hat{\alpha_n} - \hat{\alpha_{ni}})\right\}}{\alpha_n(1-\alpha_n)}
\end{aligned}$$

This can be simplified to

$$\sigma_n^N - 1 = (1-\hat{\chi}_n)(\sigma_{ni}-1) + \frac{\sum_{i\in I_n}\left\{(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}}{d\ln\omega} + (\hat{\alpha_n} - \hat{\alpha_{ni}})\right\}}{\alpha_n(1-\alpha_n)}$$

or more simply

$$\sigma_n^N = (1-\hat{\chi}_n)\sigma_{ni} + \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)}$$

To get at $\frac{d\ln\theta_{ni}}{d\ln\omega}$, we can use $Y_{ni} = Y_n P_n^{\varepsilon_n} D_{ni} P_{ni}^{-\varepsilon_n}$ and $p_{ni} = \frac{1}{\tau_{Yni}T_{Yni}}\frac{\varepsilon_n}{\varepsilon_n - 1}\hat{z}_{niY}$ to write

$$\frac{d\ln p_{ni}}{d\ln\omega} = \frac{d\ln\hat{z}_{niY}}{d\ln\omega} = 1 - \hat{\alpha_{ni}}$$

We thus have

$$\begin{aligned}
\frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)} &= \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}(-\varepsilon_n)\frac{d\ln p_{ni}}{d\ln\omega}}{\alpha_n(1-\alpha_n)} \\
&= \varepsilon_n\frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}(\alpha_{ni}-\hat{\alpha_n})}{\alpha_n(1-\alpha_n)} \\
&= \hat{\chi}_n\varepsilon_n
\end{aligned}$$

■

With this we have two results. First, if unpaid wedges do not distort any plant's capital-labor ratio ($\tau_{Kni}/\tau_{Lni} = 1$ for each i) then $\hat{\chi}_n = \chi_n$, and the formula for the aggregate elasticity is exactly the same.

Second, in Section 4.3 we describe a thought experiment in which all variation in cost shares of capital is due to unpaid wedges. To do this, we first compute the impact of unpaid wedges as follows: Define α_{ni}^* to be the i's hypothetical capital cost share is there were no unpaid wedges. Thus $\{K_{ni}^*, L_{ni}^*\} = \arg\min_{K_{ni},L_{ni}} T_{Yni}P_{ni}Y_{ni} - T_{Kni}rK_{ni} - T_{Lni}wL_{ni}$ subject to $Y_{ni} \le F_{ni}(K_{ni}, L_{ni})$ and $Y_{ni} \le Y_n P_n^{\varepsilon_n} D_{ni} P_{ni}^{-\varepsilon_n}$. This would satisfy $\frac{\alpha_{ni}}{1-\alpha_{ni}} = \left(\frac{\tau_{Kni}}{\tau_{Lni}}\right)^{-\sigma_n}\frac{\alpha_{ni}^*}{1-\alpha_{ni}^*}$ and $\frac{\hat{\alpha_{ni}}}{1-\hat{\alpha_{ni}}} = \frac{\alpha_{ni}}{1-\alpha_{ni}}\frac{\tau_{Kni}}{\tau_{Lni}}$, which together imply

$$\frac{\hat{\alpha_{ni}}}{1-\hat{\alpha_{ni}}} = \left(\frac{\alpha_{ni}}{1-\alpha_{ni}}\right)^{1-\frac{1}{\sigma_n}}\left(\frac{\alpha_{ni}^*}{1-\alpha_{ni}^*}\right)^{\frac{1}{\sigma_n}}$$

In the thought experiment, we set α_{ni}^* equal to the mean industry capital share and compute the resulting industry and aggregate elasticities. In practice, this procedure can generate extremely large and unrealistic wedges; we thus Windsorize all wedges using the 2nd and 98th percentiles.

B.5 Demand

In this section we generalize the demand system to a class of homothetic demand systems in which demand for each good is strongly separable. While this class nests Dixit-Stiglitz demand, it allows for arbitrary demand elasticities and pass through rates. An industry aggregate Y_n is defined to satisfy

$$1 = \sum_{i \in I_n} H_{ni}\left(Y_{ni}/Y_n\right) \tag{29}$$

where each H_{ni} is positive, smooth, increasing, and concave. If P_n is the ideal price index associated with Y_n, then cost minimization implies $\frac{P_{ni}}{P_n} = H'_{ni}\left(\frac{Y_{ni}}{Y_n}\right)$. Define the inverse of H'_{ni} to be $h_{ni}(\cdot) = H'^{-1}_{ni}(\cdot)$. i faces a demand curve; to find its elasticity of demand, we can differentiate:

$$d\ln Y_{ni}/Y_n = -\varepsilon_{ni}(P_{ni}/P_n)d\ln P_{ni}/P_n \tag{30}$$

where the elasticity of demand is $\varepsilon_{ni}(x) \equiv -\frac{h'_{ni}(x)x}{h_{ni}(x)}$. The optimal markup chosen by i will satisfy $\mu_{ni}(P_{ni}/P_n) = \frac{\varepsilon_{ni}(P_{ni}/P_n)}{\varepsilon_{ni}(P_{ni}/P_n)-1}$. It will be useful to define b_{ni} to be i's local relative rate of pass through: the responsiveness of P_{ni} to a change in i's marginal cost. Since $P_{ni} = \mu(P_{ni}/P_n) \times \mathrm{mc}_{ni}$, then $\frac{d\ln P_{ni}}{d\ln \mathrm{mc}_{ni}} = \frac{P_{ni}/P_n \mu'_{ni}}{\mu_{ni}}\frac{d\ln P_{ni}}{d\ln \mathrm{mc}_{ni}} + 1$, so that $b_{ni}(x) \equiv \frac{1}{1-\frac{x\mu'_{ni}(x)}{\mu_{ni}(x)}}$.

Lastly, we define $\alpha_n^P \equiv 1 - \frac{d\ln p_n}{d\ln\omega}$ to be the response of the ideal price index to a change in relative factor prices. The following claim describes the industry elasticity of substitution.

Claim 4 *Suppose that each F_{ni} exhibits constant returns to scale and the demand structure in industry n satisfies equation (29). Then the industry elasticity is*

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n + \chi_n \bar{s}_n^M \bar{\zeta}_n + \chi_n(1-\bar{s}_n^M)\bar{x}_n$$

where

$$\bar{x}_n \equiv \frac{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[(\alpha_{ni}-\alpha_n^P) - s_{ni}^M(\alpha_{ni}-\alpha^M)\right]\varepsilon_{ni}b_{ni}}{\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[(\alpha_{ni}-\alpha_n^P) - s_{ni}^M(\alpha_{ni}-\alpha^M)\right]}$$

$$\alpha_n^P = \frac{\sum_{i\in I_n} P_{ni}Y_{ni}\varepsilon_{ni}b_{ni}\left[\alpha_{ni} - s_{ni}^M(\alpha_{ni}-\alpha^M)\right]}{\sum_{i\in I_n} P_{ni}Y_{ni}\varepsilon_{ni}b_{ni}}$$

Proof. Optimal price setting implies $p_{ni} = \mu_i(p_{ni}/p_n)z_{niY}$. Taking logs and differentiating gives

$$\frac{d\ln p_{ni}/p_n}{d\ln\omega} = \frac{\mu'_i(p_{ni}/p_n)p_{ni}/p_n}{\mu_i(p_{ni}/p_n)}\frac{d\ln p_{ni}/p_n}{d\ln\omega} + \frac{d\ln z_{niY}}{d\ln\omega} - \frac{d\ln p_n}{d\ln\omega}$$

Constant returns to scale implies $\frac{d\ln z_{niY}}{d\ln\omega} = (1-\alpha_{ni})(1-s_{ni}^M) + s_{ni}^M(1-\alpha^M) = (1-\alpha_{ni}) + s_{ni}^M(\alpha_{ni} - \alpha^M)$, so this can be written as

$$\frac{d\ln p_{ni}/p_n}{d\ln\omega} = b_{ni}\left[s_{ni}^M(\alpha_{ni}-\alpha^M) - (\alpha_{ni}-\alpha_n^P)\right] \tag{31}$$

The change in output is then

$$\frac{d\ln Y_{ni}/Y_n}{d\ln\omega} = -\varepsilon_{ni}\frac{d\ln p_{ni}/p_n}{d\ln\omega} = \varepsilon_{ni}b_{ni}\left[(\alpha_{ni}-\alpha_n^P) - s_{ni}^M(\alpha_{ni}-\alpha^M)\right]$$

To get at the aggregate elasticity, we compute the following

$$\begin{aligned}
\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}}{d\ln\omega} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\frac{d\ln Y_{ni}/Y_n}{d\ln\omega} \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\varepsilon_{ni}b_{ni}\left[(\alpha_{ni}-\alpha_n^P) - s_{ni}^M(\alpha_{ni}-\alpha^M)\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\bar{x}_n\left[(\alpha_{ni}-\alpha_n^P) - s_{ni}^M(\alpha_{ni}-\alpha^M)\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\bar{x}_n\left[(\alpha_{ni}-\alpha_n^P) - \bar{s}_n^M(\alpha_{ni}-\alpha^M)\right] \\
&= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\bar{x}_n\left[(\alpha_{ni}-\alpha_n) - \bar{s}_n^M(\alpha_{ni}-\alpha_n)\right] \\
&= \alpha_n(1-\alpha_n)\chi_n(1-\bar{s}_n^M)\bar{x}_n
\end{aligned}$$

where the third equality uses the definition of $\bar{x}_n$ and the fourth uses the definition $\bar{s}_n^M$. This expression and Lemma 1 give the desired result.

It remains only to compute α_n^P. Since $\sum_{i\in I_n}\frac{P_{ni}Y_{ni}}{P_nY_n}\frac{d\ln Y_{ni}/Y_n}{d\ln\omega}=0$, we can use equation (30) and equation (31) to write

$$0 = \sum_{i\in I_n}\frac{P_{ni}Y_{ni}}{P_nY_n}\varepsilon_{ni}b_{ni}\left[s_{ni}^M(\alpha_{ni}-\alpha^M) - (\alpha_{ni}-\alpha_n^P)\right]$$

which simplifies to

$$\alpha_n^P = \frac{\sum_{i\in I_n}P_{ni}Y_{ni}\varepsilon_{ni}b_{ni}\left[\alpha_{ni} - s_{ni}^M(\alpha_{ni}-\alpha^M)\right]}{\sum_{i\in I_n}P_{ni}Y_{ni}\varepsilon_{ni}b_{ni}}$$

■

B.6 Non-Homothetic Production

This section analyzes how the industry elasticity of substitution is altered if production is non-homothetic. This requires a more careful definition of the elasticities of substitution. A change in factor prices will have a direct effect on a plant's choice of capital-labor ratio, and may have an indirect impact if the change in factor prices alters a plant's scale. We pursue an approach similar to Joan Robinson: we define a plant's elasticity of substitution to be how a change in relative factor prices alters the plant's capital-labor ratio holding output fixed. Similarly, an industry's elasticity of substitution is the response of the industry's capital labor ratio to a change in relative factor prices holding fixed the industry aggregate, Y_n.

We first characterize the plant-level elasticity of substitution, and then derive an expression for the industry level elasticity. In the interest of space, we restrict attention to the case in which plants do not use materials.

Just as $1-\alpha_{ni}\left(=\frac{z_{ni\omega}\omega}{z_{ni}}\right)$ is the labor share of i's cost, we define $\tilde{\alpha}_{ni}$ so that $1-\tilde{\alpha}_{ni}=\frac{z_{niY\omega}\omega}{z_{niY}}$, the labor share of i's marginal cost.

Since $1-\alpha_{ni} = \frac{z_{ni\omega}(Y_{ni},\omega)\omega}{z_{ni}(Y_{ni},\omega)}$, we have

$$d\ln(1-\alpha_{ni}) = \frac{z_{ni\omega Y}Y_{ni}}{z_{ni\omega}}d\ln Y_{ni} + \frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}d\ln\omega + d\ln\omega - \frac{z_{niY}Y_{ni}}{z_{ni}}d\ln Y_{ni} - \frac{z_{ni\omega}\omega}{z_{ni}}$$

Since $\frac{z_{ni\omega Y}Y_{ni}}{z_{ni\omega}} = \frac{z_{ni\omega Y}\omega}{z_{niY}}\frac{z_{niY}Y_{ni}}{z_{ni}}$, this can be arranged as

$$d\ln(1-\alpha_{ni}) = \left(\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}} + 1 - (1-\alpha_{ni})\right)d\ln\omega + \left(\frac{1-\tilde{\alpha}_{ni}}{1-\alpha_{ni}} - 1\right)\frac{1}{\gamma_{ni}}d\ln Y_{ni}$$

Using $d\ln\frac{\alpha_{ni}}{1-\alpha_{ni}} = -\frac{1}{\alpha_{ni}}d\ln(1-\alpha_{ni})$, we have

$$d\ln\frac{\alpha_{ni}}{1-\alpha_{ni}} = \left(-\frac{1}{\alpha_{ni}}\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}} - 1\right)d\ln\omega + \frac{\tilde{\alpha}_{ni}-\alpha_{ni}}{\alpha_{ni}(1-\alpha_{ni})}\frac{1}{\gamma_{ni}}d\ln Y_{ni}$$

By definition, $\sigma_{ni}-1$ is the change in $\frac{\alpha_{ni}}{1-\alpha_{ni}}$ holding Y_{ni} fixed. The plant level elasticity of substitution is

$$\sigma_{ni} = -\frac{1}{\alpha_{ni}}\frac{z_{ni\omega\omega}\omega}{z_{ni\omega}}$$

and

$$d\ln\frac{\alpha_{ni}}{1-\alpha_{ni}} = (\sigma_{ni}-1)d\ln\omega + \frac{\tilde{\alpha}_{ni}-\alpha_{ni}}{\alpha_{ni}(1-\alpha_{ni})}\frac{1}{\gamma_{ni}}d\ln Y_{ni} \tag{32}$$

Claim 5 *The industry elasticity is*

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n + \tilde{\chi}_n\bar{x}_n$$

where χ_n and $\bar{\sigma}_n$ are defined as in Lemma 1 and

$$\tilde{\chi}_n \equiv \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_n)(\tilde{\alpha}_{ni}-\alpha_n^P)\theta_{ni}}{\alpha_n(1-\alpha_n)}$$

$$\bar{x}_n \equiv \frac{\sum_{i\in I_n}(\tilde{\alpha}_{ni}-\alpha_n)(\tilde{\alpha}_{ni}-\alpha_n^P)\theta_{ni}\frac{\varepsilon_n/\gamma_{ni}}{1+\varepsilon_n\frac{z_{niYY}Y_{ni}}{z_{niY}}}}{\sum_{i\in I_n}(\tilde{\alpha}_{ni}-\alpha_n)(\tilde{\alpha}_{ni}-\alpha_n^P)\theta_{ni}}$$

and α_n^P is defined to satisfy $1-\alpha_n^P = \frac{d\ln p_n}{d\ln\omega}$.

Proof. Following the same logic as in the benchmark, we have

$$d\ln\frac{\alpha_n}{1-\alpha_n} = \sum_{i\in I_n}\frac{\alpha_{ni}(1-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)}d\ln\frac{\alpha_{ni}}{1-\alpha_{ni}} + \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)}{\alpha_n(1-\alpha_n)}\theta_{ni}d\ln\theta_{ni}$$

Using equation (32), this becomes

$$\begin{aligned}
d\ln\frac{\alpha_n}{1-\alpha_n} &= \sum_{i\in I_n}\frac{\alpha_{ni}(1-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)}(\sigma_{ni}-1)d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1}{\gamma_{ni}}d\ln Y_{ni} + \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)}{\alpha_n(1-\alpha_n)}\theta_{ni}d\ln\theta_{ni} \\
&= (1-\chi_n)(\bar{\sigma}_n-1)d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_{ni})\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1}{\gamma_{ni}}d\ln Y_{ni} + \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)}{\alpha_n(1-\alpha_n)}\theta_{ni}d\ln\theta_{ni}
\end{aligned}$$

where the second line used the definitions of $\bar{\sigma}_n$ and χ_n. Since $\theta_{ni} = z_{ni}/\sum_{j\in I_n} z_{nj}$, we have

$$\begin{aligned}\sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni} d\ln\theta_{ni} &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{z_{niY}Y_{ni}}{z_{ni}}d\ln Y + \frac{z_{ni\omega}\omega}{z_{ni}}d\ln\omega - d\ln\sum_{j\in I_n} z_{nj}\right] \\ &= \sum_{i\in I_n}(\alpha_{ni}-\alpha_n)\theta_{ni}\left[\frac{1}{\gamma_{ni}}d\ln Y + (1-\alpha_{ni})d\ln\omega\right]\end{aligned}$$

Plugging this in and combining coefficients gives

$$d\ln\frac{\alpha_n}{1-\alpha_n} = (1-\chi_n)(\bar{\sigma}_n - 1)d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_n)\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1}{\gamma_{ni}}d\ln Y_{ni} + \sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)}{\alpha_n(1-\alpha_n)}\theta_{ni}(1-\alpha_{ni})d\ln\omega$$

One can easily verify that $\sum_{i\in I_n}\frac{(\alpha_{ni}-\alpha_n)(\alpha_{ni}-1)}{\alpha_n(1-\alpha_n)}\theta_{ni} = \chi_n$. This and $d\ln\frac{\alpha_n}{1-\alpha_n} = d\ln K_n/L_n - d\ln\omega$ imply

$$d\ln K_n/L_n = (1-\chi_n)\bar{\sigma}_n d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_n)\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1}{\gamma_{ni}}d\ln Y_{ni} \tag{33}$$

Finally we need to address the changes in scale. i's price is $p_{ni} = \frac{\varepsilon_n}{\varepsilon_n - 1}z_{niY}$, so the change in i's price is

$$\begin{aligned}d\ln p_{ni} &= \frac{z_{niYY}Y_{ni}}{z_{niY}}d\ln Y_{ni} + \frac{z_{niY\omega}\omega}{z_{niY}}d\ln\omega \\ &= \frac{z_{niYY}Y_{ni}}{z_{niY}}d\ln Y_{ni} + (1-\tilde{\alpha}_{ni})d\ln\omega\end{aligned}$$

If the change in the industry price index satisfies $d\ln p_n = (1-\alpha_n^P)d\ln\omega$, then the change in output is

$$\begin{aligned}d\ln Y_{ni} &= -\varepsilon_n d\ln\frac{p_{ni}}{p_n} + d\ln Y_n \\ &= -\varepsilon_n\left(\frac{z_{niYY}Y_{ni}}{z_{niY}}d\ln Y_{ni} + (1-\tilde{\alpha}_{ni})d\ln\omega - (1-\alpha_n^P)d\ln\omega\right) + d\ln Y_n \\ &= \frac{\varepsilon_n(\tilde{\alpha}_{ni}-\alpha_n^P)d\ln\omega + d\ln Y_n}{1+\varepsilon_n\frac{z_{niYY}Y_{ni}}{z_{niY}}}\end{aligned}$$

Using the definition of $\bar{x}_n$ and $\tilde{\chi}_n$, we therefore have that

$$\begin{aligned}d\ln K_n/L_n &= (1-\chi_n)\bar{\sigma}_n d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_n)\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1}{\gamma_{ni}}\frac{\varepsilon_n(\tilde{\alpha}_{ni}-\alpha_n^P)d\ln\omega + d\ln Y_n}{1+\varepsilon_n\frac{z_{niYY}Y_{ni}}{z_{niY}}} \\ &= (1-\chi_n)\bar{\sigma}_n d\ln\omega + \tilde{\chi}_n\bar{x}_n d\ln\omega + \sum_{i\in I_n}\frac{(\tilde{\alpha}_{ni}-\alpha_n)\theta_{ni}}{\alpha_n(1-\alpha_n)}\frac{1/\gamma_{ni}}{1+\varepsilon_n\frac{z_{niYY}Y_{ni}}{z_{niY}}}d\ln Y_n\end{aligned}$$

Since σ_n^N is defined to be the change in K_n/L_n in response to a change in ω holding fixed Y_n, we have

$$\sigma_n^N = (1-\chi_n)\bar{\sigma}_n + \tilde{\chi}_n\bar{x}_n$$

C Data Notes

C.1 Census of Manufacturers

We use the 1987 and 1997 Census of Manufactures to estimate plant elasticities of substitution and demand. We remove all Administrative Record plants because these plants do not have data on output or capital. We also eliminate a set of outliers and missing values from the dataset. We first remove all plants born in the given Census year, as well as a small set of plants with missing age data. We then remove plants with zero or missing data on the following variables: average revenue product of capital, average revenue product of labor, capital share, capital labor ratio, and plant level wage. We also remove plants above the 99.5th percentile or below the 0.5th percentile of their 4 digit SIC industry on these variables to remove plants with potential data problems.

For capital costs, we multiply capital stock measures by rental rates of capital. In the 1987 Census, the Census asked plants to report the book value of structures capital separately from equipment capital. Thus, we construct the capital stock for structures capital separately from equipment capital for 1987. Because the book value reported in the Census is a historical gross cost measure (although it accounts for capital retirements), we multiply the book value of capital by a current net cost to historical gross cost deflator based upon estimates of the current net value of capital and historic gross value of capital constructed by the Bureau of Economic Analysis at the 2 digit SIC level. Because this deflator is not base 1987, we then use investment deflators to convert each capital stock to 1987 dollars.

In the 1997 Census, the Census only asked plants to report the total value of capital. We construct capital deflators and rental rates for both structures and equipment capital using the same procedure as 1987. We then average both the capital deflator and rental rate of capital for structures and equipment capital, weighting each type of capital by its share of overall capital based upon data for the plant's 4 digit SIC industry from the NBER Productivity Database.

C.2 Local Wages

We construct measures of the local wage in order to estimate the elasticity of substitution across plants, using two different datasets to measure the local area wage. The primary dataset that we use in the Census 5 percent samples of Americans. The Population Censuses have data on both wages and MSA geographic location for a large sample of workers.

To obtain the local wage, we first calculate the individual wage for prime age men (with age between 25 and 55) who are employed in the private sector as workers earning a wage or salary. We calculate the wage as an hourly wage, defined as total yearly wage and salary income divided by total hours worked. We measure total hours worked as weeks worked per year multiplied by hours worked per week. We remove all individuals with zero or missing income or zero total hours worked. For 1990, incomes above the Census top code of $140,000 are set to the state median of wage and salary income above the top code. For 2000, incomes above the Census top code of $175,000 are set to the state mean of wage and salary income above the top code.

Before calculating local area wages, we adjust measures of local wages for differences in worker characteristics through regressions with the individual log wage as a dependent variable. We include education through a set of dummy variables based upon the worker's maximum educational attainment, which include four categories: college, some college, high school degree, and high school dropouts. We define experience as the individual's age minus an initial age of working that depends upon their education status, and include a quartic in experience in the regression. We also have data on the race of workers and so include three race categories of white, black, and other. We include six occupational categories: Managerial and Professional; Technical, Sales, and Administrative; Service, Farming, Forestry, and Fishing; Precision Production, Craft, and Repairers; and Operatives and Laborers. Finally, we include thirteen industrial categories: Agriculture, Forestry, and Fisheries; Mining; Construction; Manufacturing; Transportation, Communications and Other Public Utilities; Wholesale Trade; Retail Trade; Finance, Insurance, and Real Estate; Business and Retail

Services; Personal Services; Entertainment and Recreation Services; Professional and Related Services; and Public Administration. We then calculate the local area wage as the MSA average of residual wages from a regression that includes all of these characteristics, with separate regressions by year. Because the Economic Census is conducted in different years from the Population Censuses, we match the 1987 Census of Manufactures to wages from the 1990 Population Census, and the 1997 Census of Manufactures to wages from the 2000 Population Census.

The second dataset that we use for robustness checks is the Longitudinal Business Database, which contains data on payroll and employment for all US establishments. We construct the establishment wage as total payroll divided by total employment. We measure the local wage as the mean log wage at the county level. We match the 1987 Longitudinal Business Database to the 1987 Census of Manufactures and the 1997 Longitudinal Business Database to the 1997 Census of Manufactures.

C.3 Instruments

We use labor demand instruments for the local wage for robustness checks on our estimates of the elasticity of substitution, based upon the differential impact of national level shocks to industry employment across locations. Positive national shocks to an industry should increase labor demand and wages, more in areas with high concentrations of that industry. Formally, the predicted growth rate in employment for a given location is the sum across industries of the product of the local employment share of this industry and the 10 year change in national level employment for that industry. We use the Longitudinal Business Database, which contains all US establishments, to construct these instruments.

The implicit assumption here is that changes in industry shares at the national level are independent of local manufacturing plant productivity. To help ensure that this assumption holds, we exclude manufacturing industries from the labor demand instrument. We calculate the instrument defining locations by MSAs and industries at the SIC 4 digit level. For 1987, we use the instrument from 1976-1986 because the SIC 4 digit industry definitions change significantly from 1977 to 1987.

C.4 Annual Survey of Manufactures

The Annual Survey of Manufactures tracks about 50,000 plants over five year panel rotations that are more heavily weighted towards large plants. We use the ASM to calculate the heterogeneity indices and materials shares. The ASM has data on plant investment over time as well as book values of the stock of capital, which we use to construct perpetual inventory measures of capital.

We also take into account retirements of the capital stock until 1987, as data on retirements of capital stock are available from 1977-1987 excepting 1986. For 1973-1976 and 1986 we can calculate an imputed value for retirements as end of year capital subtracted from beginning year capital and yearly investment; we lower investment if this value is negative. Plants retire their capital stock at a rate of about 4 percent a year, which is concentrated in a few plants retiring a lot of capital stock. Since firms retiring capital deduct the retirement values from their book value, the book value incorporates depreciation from retirements.

We calculate perpetual inventory measures of capital through the following capital accumulation equation, as in Caballero et al. (1995):

$$K_t = (1 - \delta^a)K_{t-1} + I_t - R_t$$

where K_t is period t capital stock, I_t is period t investment, R_t is period t retirements, and δ^a is the in use depreciation rate. We build separate capital stocks for structures and equipment capital. To calculate the in use depreciation rate δ^a, we first calculate δ^r the average yearly rate of capital retirements (total retired capital stock divided by beginning gross capital stock) across plants from 1977 to 1985 by 2 digit SIC industry. We then initially define the in use depreciation rate as:

$$\delta^a = \delta - \delta^r$$

where δ is the overall 2 digit SIC depreciation rate calculated by the BLS minus this yearly retirement rate.

We account for retirements by building a set of capital vintages for each year that the plant exists in the dataset. Retirements are taken out of the gross capital stock of the earliest vintages of capital, as we assume FIFO retirement of capital. We initialize capital stock by the initial sample year book value, so for the first year that the plant exists in the dataset, capital is set to book value of capital. We deflate this book value by a net current cost to gross historical cost deflator. In subsequent years, each vintage is investment deflated through the investment deflator. Real investment is added to capital, and in use depreciation subtracted from capital. After this process, we recalculate the retirement depreciation rate as capital retired net of in use depreciation divided by net overall capital stock, and then recalculate all of the capital vintages to construct an overall capital measure.

After 1987, retirements are no longer recorded so we calculate perpetual inventory measures of capital without retirements, as in the following capital accumulation equation:

$$K_t = (1 - \delta)K_{t-1} + I_t$$

where δ is the overall depreciation rate. After 1992, the Census no longer records book values for structures and equipment separately, although they do record investment separately by capital type. When we only have a total book value of capital for a plant, we use earlier data from the plant on the share of equipment capital to form separate capital stocks for structures and equipment. If no earlier data from the plant is available, we use the share of equipment capital for the 4 digit industry from the NBER productivity database.

The ASM plant samples also have data on the value of non monetary compensation given to employees, such as health care or retirement benefits, which we use to better measure payments to labor.

C.5 Rental Rates

We define the rental rate using the external real rate of return specification of Harper et al. (1989). The rental rate for industry n is defined as:

$$R_{i,t} = T_{i,t}(p_{i,t-1}r_{i,t} + \delta_{i,t}p_{i,t})$$

where $r_{i,t}$ is a constant external real rate of return of 3.5 percent, $p_{i,t}$ is the price index for capital in that industry, $\delta_{i,t}$ is the depreciation rate for that industry, and $T_{i,t}$ is the effective rate of capital taxation. We calculate $T_{i,t}$ following Harper et al. (1989) as:

$$T_{i,t} = \frac{1 - u_t z_{i,t} - k_{i,t}}{1 - u_t}$$

where $z_{i,t}$ is the present value of depreciation deductions for tax purposes on a dollar's investment in capital type i over the lifetime of the investment, $k_{i,t}$ is the effective rate of the investment tax credit, and u_{it} is the effective corporate income tax rate. We obtained $z_{i,t}$, u_{it}, and $k_{i,t}$ from Dale Jorgenson at the asset year level; we then used a set of capital flow tables at the asset-industry level to convert these to the industry level.

To calculate depreciation rates $\delta_{i,t}$, we take depreciation rates from NIPA at the asset level and use the capital flow tables to convert them to the industry level. Our primary source of prices of capital $p_{i,t}$ are from NIPA, which calculates separate price indices for structures and equipment capital. As an alternative, we also develop a set of rental rates based upon the investment price series of Cummins and Violante (2002).

The capital flow tables and investment price series depend upon the industry definition; because the US switches from SIC basis to NAICS basis during this period, we construct separate rental price series for SIC 2 digit industries and NAICS 3 digit industries. Finally, when we examine the aggregate we have to aggregate all of the rental price series; we do so by calculating Tornqvist indices between equipment and structures capital for each industry, and then a Tornqvist index across rental rates for each industry. The Tornqvist indices allow for the share of equipment capital in industry capital and for the share of different industries in manufacturing capital to change over this period.

C.6 Homogenous Product Industries

We follow a similar process to Foster et al. (2008) in constructing data on homogenous product industries for robustness checks on the elasticity of demand. We use six homogenous products: Boxes, Bread, Coffee, Concrete, Processed Ice and Plywood.[42] All of the products are defined as in Foster et al. (2008). We use data from 1987-1997 as capital data was imputed before 1987 for non-ASM plants, although we do not use data for 1992 for Processed Ice (because of data errors), 1987 for Boxes (because of a product definition change), and 1997 for Concrete (because quantity data was not recorded). We remove Census balancing codes imputed by the Census to make product level data add up to overall revenue data in cases where we can identify them. We also remove receipts for contract work, miscellaneous receipts, resales of products, and products with negative values.

We then remove all plants for which the product's share of plant revenue (measured after removing the balancing codes and other items mentioned above) is less than 50 percent. For each product, we have measures of both total quantity produced and revenue, which allows me to calculate product price as revenue over quantity. We delete all plants for which the ratio of product price to median product price is between .999 and 1.001, as these plants likely have quantity data imputed by the Census. We also remove plants with prices greater than ten times the median price or less than one-tenth the median price as potential mismeasured outliers.

C.7 Cross Country Data

We obtain plant-level data for Chile, Colombia, and India from national plant-level manufacturing censuses. The Chilean data spans 1986 to 1996 with about 5,000 plants per year, the Colombian data from 1981 to 1991 with about 7,000 plants per year, and the Indian data from 2000 to 2003 with about 30,000 plants per year.[43] The Chilean and Colombian data cover all manufacturing plants with at least ten employees, while the Indian data are a sample of all plants with at least ten employees (twenty if without power), with plants with at least one hundred workers sampled with certainty. We define industries at a similar level to two digit US SIC; for Chile and Colombia this is at the three digit ISIC level, and for India at the two digit NIC level.

Capital costs are the most involved variable to construct. For each country, a capital stock is constructed for each type of capital. Capital services is the sum of the stock of each type multiplied by its rental rate plus rental payments. To account for utilization (and especially entry and exit), we multiply capital services by the fraction of the year the plant was open. The capital rental rate is composed of the real interest rate R, depreciation rate δ for that type of capital, and effective corporate tax rate τ:

$$r = \frac{R+\delta}{1-\tau}$$

For corporate tax rates, we use the one year effective tax rate collected by Djankov et al. (2010).Djankov et al. (2010) derive effective tax rates for fiscal year 2004 by asking a major accounting firm to calculate the tax rate for the same fictitious corporation in 85 countries.

Across countries, there are some differences in the construction of capital stocks and depreciation rates. For Chile, we use capital stocks constructed by Greenstreet (2007). Greenstreet (2007) constructed capital stocks for each type of capital using a permanent inventory type procedure. We use his depreciation rates of 5 percent for buildings, 10 percent for equipment, and 20 percent for vehicles.

For Colombia and India, we construct measures of capital services. To construct capital for Colombia, we broadly follow the perpetual inventory procedure of Tybout and Roberts (1996). Because the Indian data is not panel, we use book values of capital for each type of capital. For both Colombia and India, we match the depreciation rates we calculate for US industries to the equivalent industries in Colombia and India for structures and equipment, while for transportation, we follow Greenstreet (2007) and set the depreciation

[42] Foster et al. (2008) examine 5 additional products: Carbon Black, Flooring, Gasoline, Block Ice, and Sugar; small samples in the years we study preclude this analysis.

[43] We have data from Chile going back until 1979, but we only use the later years to avoid the Chilean financial crisis in the early 1980s.

rate to 0.20.[44]

We base the real interest rate on private sector lending rates reported in the IMF Financial Statistics. For Colombia, we have capital deflators over time and so construct separate real interest rates for each type of capital by deducting the realized inflation rates for each type of capital from the lending rate. For India, we do not have investment deflators and so use the GDP deflator.

We then use the average real interest rate over our sample period for the rental rates. For labor costs, we use the available wages and benefits data for each country.

To construct rental rates of capital for our policy experiments, we require real interest rates, corporate tax rates, and depreciation rates for each country at the same point in time. For the real interest rate, we adjust the nominal private sector lending rate for each country from the IMF International Financial Statistics for inflation, and then average from 1992 to 2011.[45] For corporate tax rates, we again use the one year effective tax rate collected by Djankov et al. (2010). Finally, we set the depreciation rate to 9.46 percent based upon US manufacturing data.

D Additional Empirical Results

D.1 Local Content of Materials

In our baseline estimates of the elasticity of substitution between materials and non-materials, ζ, we assume that the local wage does not affect the materials price the plant faces. As a robustness check, we examine how sensitive our estimates are to correlation between materials prices and local wages due to local content of materials. The local wage would affect labor costs for locally sourced materials. We use the 1993 Commodity Flow Survey to construct the local content of shipments for every industry included in the survey, defining local as a shipment within 100 miles of the originating factory. We then use the 1992 Input-Output tables to construct the average local content of materials for every manufacturing industry. Assuming that every input industry has the same materials and labor shares and fraction of local content of materials, the elasticity of the materials price with respect to the wage is:

$$\frac{d\log q_i}{d\log w} = (1-\alpha_n)\frac{1-s_n^M l_c}{1-(1-s_n^M)l_c}$$

where l_c is the measure of local content.

We therefore estimate ζ using the regression

$$\log\frac{rK_i+wL_i}{qM_i} = (1-\zeta)\frac{(1-\alpha_i)}{(1-\alpha_n)\frac{1-s_n^M l_c}{1-(1-s_n^M)l_c}}(\log w_i) + CONTROLS + \epsilon_i$$

As we report in the text, we find only slightly lower estimates in estimated elasticities after accounting for the local content of materials.

D.2 Cross Industry Demand Elasticity

The cross industry elasticity of demand characterizes how industry level demand responds to a change in the overall industry price level. To estimate this elasticity, we use panel data on quantity and price at the industry level from the NBER productivity database from 1962 to 2009.

Since least squares estimates conflate demand and supply, we have to instrument for price using supply side instruments that capture industry productivity. The two instruments that we examine is the average product of labor, defined as the amount of output produced per worker, and the average real cost per unit

[44] The US depreciation rates are based on NIPA data on depreciation rates of assets; we then use asset-industry capital tables to construct depreciation rates for structures and equipment for each industry.

[45] We employ a discrete time correction as some countries have high inflation rates, so $R = \frac{i_t-\pi_t}{1+\pi_t}$ for lending rate i_t and inflation rate π_t. We use the change in the GDP deflator for inflation.

of output produced, which is the appropriate measure of industry productivity in our model. We thus have the following regression specification:

$$\log q_{n,t} = -\eta \log p_{n,t} + \alpha_n + \beta_t + CONTROLS + \varepsilon_n$$

where $q_{n,t}$ is quantity produced for industry n in period t, $p_{n,t}$ is the price for industry n in period t, α_n are a set of industry fixed effects, and β_t are a set of time fixed effects.

We then examine the cross industry demand elasticity, defining industry at both the four digit and two digit SIC levels. We have 459 four digit industries and 20 two digit industries.[46] For each industry definition, we develop specifications with extra sets of controls to account for potential trends over time that could be correlated with changes in prices. In the four digit specifications, these extra controls include either 2 digit industry-year fixed effects, or 4 digit industry linear trends. In the two digit specifications, these extra controls include 2 digit industry linear trends.

Table VI below contains these estimates, as well as the OLS estimate. As would be expected from simultaneity bias, OLS estimates are lower in magnitude than IV estimates. The IV estimates using four digit industries range between 1.2 and 2.2 and are slightly above estimates using two digit industries. This pattern is consistent with two digit industry varieties being less substitutable than four digit industry varieties.

The two digit industry IV estimates range from 0.75 to 1.15, with three of the four estimates close to one. Because we define industries in our aggregation analysis at the two digit level, the two digit industry estimates are more appropriate. We thus set the cross industry demand elasticity to one. Our results are not extremely sensitive to this elasticity; increasing the elasticity from 1 to 1.5 would increase the US aggregate elasticity by about 0.01.

Table VI Cross Industry Elasticity of Demand for the Manufacturing Sector

	Industry Definition:				
Instrument		Four Digit			Two Digit
None	0.99	1.06	0.57	0.91	0.37
	(0.02)	*(0.01)*	*(0.02)*	*(0.03)*	*(0.05)*
APL	1.30	1.28	2.12	1.14	1.05
	(0.01)	*(0.01)*	*(0.03)*	*(0.04)*	*(0.06)*
Avg Cost	1.19	1.22	1.58	1.04	0.77
	(0.01)	*(0.01)*	*(0.02)*	*(0.03)*	*(0.05)*
Industry-Year Controls	None	Two Digit FE	Four Digit Trends	None	Two Digit Trends

Note: Standard errors are in parentheses. The first row contains coefficients from OLS regressions, while the second and third row are IV regressions with either the average product of labor or average real cost per unit produced as instruments. The first three columns are on four digit SIC industries; all regressions contain four digit SIC industry and year fixed effects. The second column also includes two digit industry-year fixed effects and the third column also includes four digit industry linear time trends. The last two columns are on two digit SIC industries; all regressions contain two digit SIC industry and year fixed effects. The last column also includes two digit industry linear time trends.

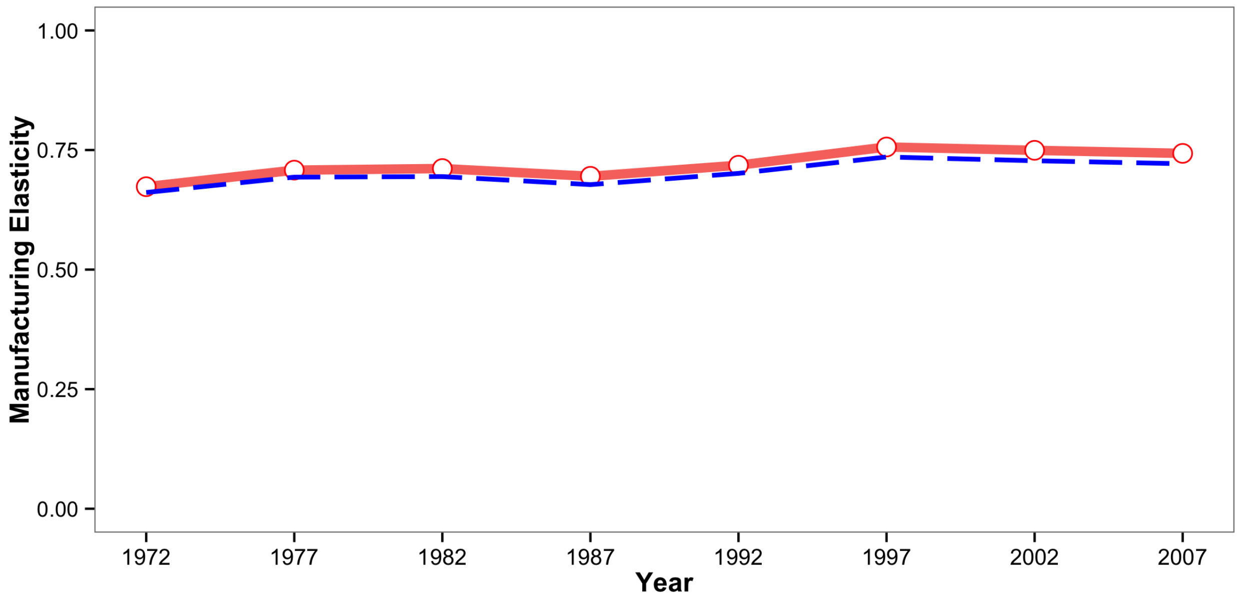
D.3 Aggregate Elasticity Over Time

Our baseline approach to examining the change in the aggregate elasticity over time fixes plant demand and production elasticities at their 1987 values. In this section, we show how these estimates change if we

[46]Since the underlying data is at the four digit industry level, we develop two digit SIC prices and quantities using a Fisher ideal index with base year 1987. We also exclude eight 4 digit industries which disappear because they are excluded after the Census shifts to NAICS basis manufacturing, the most prominent of which is Newspaper Publishing.

instead use the production and demand elasticities from 1997. Figure 11 depicts the aggregate elasticity of substitution over time both using the 1997 micro elasticities, in dashed blue, and using the 1987 elasticities in solid red. The 1997 elasticities lower all of the estimates by about 0.01. Almost all of this change is due to the elasticity of materials to non-materials falling from 0.90 to 0.67; just changing the plant demand elasticities and capital-labor substitution elasticities has almost no effect on the aggregate estimates.

Figure 11 Aggregate Elasticity of Substitution Across Time Using 1997 Production and Demand Elasticities



Note: The figure displays the manufacturing level elasticity of substitution by Census year from 1972-2007. The red line is based on the 1987 plant demand and substitution elasticities, and the blue line is based on the 1997 elasticities.

D.4 Labor Share Decomposition

D.4.1 Derivation of Decomposition

This section describes the how the theory is used to execute the decomposition of equation (12). Labor's share of value added is $s^{v,L} = \sum_n v_n s_n^{vL}$, where $s_n^{v,L}$ is labor's share of added in industry n and v_n is industry n's share of total value added. Changes in labor's share of value added can come from changes in labor's share within industries or changes between industries:

$$\begin{aligned} ds^{v,L} &= \sum_n s_n^{v,L} dv_n + \sum_n v_n ds_n^L \\ &= \sum_n \left(s_n^{v,L} - s^{v,L}\right) dv_n + \sum_n v_n ds_n^{v,L} \end{aligned}$$

We can decompose each into the components that responded to price changes, a within-industry residual, and a between-industry residual.

$$ds^{v,L} = \frac{\partial s^{v,L}}{\partial w/r} d\ln w/r + \sum_n \left(s_n^{v,L} - s^{v,L}\right) \left(dv_n - \frac{\partial v_n}{\partial \ln w/r}\right) d\ln w/r + \sum_n v_n \left(ds_n^{v,L} \frac{\partial s_n^{v,L}}{\partial \ln w/r} d\ln w/r\right)$$

where $\frac{\partial s^{v,L}}{\partial \ln w/r} = \sum_n \left(s_n^{v,L} - s^{v,L}\right) \frac{\partial v_n}{\partial \ln w/r} + \sum_n v_n \frac{\partial s_n^{v,L}}{\partial \ln w/r}$.

We now derive expressions for $\frac{\partial s_n^{v,L}}{\partial w/r}$ and $\frac{\partial v_n}{\partial w/r}$. To get at these we will use

$$\begin{aligned}\frac{\partial \ln(1-\alpha_n)}{\partial \ln w/r} &= \alpha(1-\sigma_n)\\ \frac{\partial \ln\left(1-s_n^M\right)}{\partial \ln \omega} &= s_n^M\left(1-\zeta_n^N\right)\left(\alpha^M-\alpha_n\right)\end{aligned}$$

Define $\mu_n = \frac{\varepsilon_n}{\varepsilon_n - 1}$. Since $s_n^{v,L} = (1-\alpha_n)\frac{rK_n+wL_n}{VA_n} = (1-\alpha_n)\frac{1-s_n^M}{1-s_n^M+\mu_n}$. We then have

$$\begin{aligned}\frac{\partial s_n^{v,L}}{\partial w/r} &= s_n^{v,L}\left\{\frac{\partial \ln 1-\alpha_n}{\partial \ln w/r} + \frac{\mu_n}{1-s_n^M+\mu_n}\frac{\partial \ln\left(1-s_n^M\right)}{\partial \ln w/r}\right\}\\ &= s_n^{v,L}\left\{\alpha_n\left(1-\sigma_n^N\right) + \frac{\mu_n}{1-s_n^M+\mu_n}s_n^M\left(1-\zeta_n^N\right)\left(\alpha^M-\alpha_n\right)\right\}\end{aligned}$$

To get at the between terms, note that

$$v_n = \frac{VA_n}{VA} = \left(1-s_n^M+\mu_n\right)\left(\frac{rK_n+wL_n+qM_n}{rK+wL+qM}\right)\left(\frac{rK+wL+qM}{VA}\right)$$

Since $\sum_n v_n(s_n^{v,L}-s^{v,L}) = 0$, we have

$$\sum_n v_n(s_n^{v,L}-s^{v,L})\frac{\partial \ln v_n}{\partial \ln w/r} = \sum_n v_n\left(s_n^{v,L}-s^{v,L}\right)\left[\frac{\partial \ln\left(1-s_n^M+\mu_n\right)}{\partial \ln w/r} + \frac{\partial \ln \frac{rK_n+wL_n+qM_n}{rK+wL+qM}}{\partial \ln w/r}\right]$$

Since $\eta = 1$, $\frac{\partial \ln \frac{rK_n+wL_n+qM_n}{rK+wL+qM}}{\partial \ln w/r} = 0$. Thus

$$\begin{aligned}\sum_n v_n(s_n^{v,L}-s^{v,L})\frac{\partial \ln v_n}{\partial \ln w/r} &= \sum_n v_n\left(s_n^{v,L}-s^{v,L}\right)\frac{1-s_n^M}{1-s_n^M+\mu_n}\frac{\partial \ln\left(1-s_n^M\right)}{\partial \ln w/r}\\ &= \sum_n v_n\left(s_n^{v,L}-s^{v,L}\right)\frac{\left(1-s_n^M\right)s_n^M}{1-s_n^M+\mu_n}\left(1-\zeta_n^N\right)\left(\alpha^M-\alpha_n\right)\end{aligned}$$

D.4.2 Profit Share and Intangible Capital

In our benchmark decomposition, we assume that factor prices do not impact $s^{v,\pi}$. This assumption implicitly views changes in $s^{v,\pi}$ as stemming from shifts in preferences that impact the elasticity of substitution or from changes in the share of materials. This section takes an alternative view: that $s^{v,\pi}$ reflects a return to intangible capital. If a plant's use of intangible capital responds to factor prices, then $s^{v,\pi}$ would respond to factor prices as well. We therefore develop a decomposition of the labor share under the assumption that the intangible capital/labor ratio responds to r/w in the same way as the ratio of fixed capital to labor.

$$d\ln\frac{s^{v,L}}{s^{v,K}+s^{v,\pi}} = (\sigma-1)d\ln\frac{r}{w} + d\ln\phi$$

The change in the labor share can be written as

$$ds^{v,L} = s^{v,L}(1-s^{v,L})d\ln\frac{s^{v,L}}{1-s^{v,L}} = s^{v,L}(1-s^{v,L})(\sigma-1)d\ln\frac{r}{w} + s^{v,L}(1-s^{v,L})d\ln\phi$$

Table VII displays our decomposition under the intangible capital framework. Compared to our baseline results, the contribution from factor prices to the rise in the labor share rises to about 0.12 percentage points per year from 0.07 percentage points per year in the 1970-1999 period. In the 2000-2010 period, this rise is smaller, going to 0.07 percentage points per year to 0.05 percentage points per year. The magnitude of contribution from biased technical change also increases to offset the contribution from factor prices.

Table VII Contributions to Labor Share Change

		Contributions from:	
Period	Labor Share Change	Factor Prices	Capital
1970-1999	-0.25%	0.12%	-0.38%
2000-2010	-0.79%	0.07%	-0.87%

Note: All changes are annualized. The capital contribution includes the contribution from fixed and intangible capital.

D.4.3 Labor Share from Production Data

Our benchmark analysis decomposed labor's share of income as measured in the national accounts. This data is built from manufacturing firms. Alternatively, we could analyze the changes in labor share as measured from production data built from manufacturing plants. We will briefly describe the advantages of each and why the analysis based on national accounts is our preferred measure.

The national accounts is built from firm data, so it includes all establishments (including non-manufacturing establishments) of manufacturing firms. This data contains measures of overall labor compensation.

The production data from the NBER CES production database is built from the same manufacturing plant database that we used to compute the aggregate elasticity. Because the aggregate production data does not include benefits, in each year we adjust the payments to labor by the ratio of total compensation to wages and salaries for manufacturing from NIPA.

We prefer using the labor share from the national accounts for two reasons. First, it makes our study comparable to the rest of the literature that has studied the labor share. Second, the production data only includes expenses incurred at the plant level, such as energy and materials costs. It does not include expenses such as advertising, research and development not conducted at the plant, and all expenses at the corporate headquarters The absence of these expense means that value added, and hence our residual "profit", are both overstated and may have different trends over time.

Nevertheless, Table VIII displays the change in the labor share and its components under two alternatives. First, we perform the same analysis as in the text, decomposing the change in the labor share of value added as measured in the production data. Second, we decompose the change in labors share of the total expenditure on capital and labor, $d\frac{s_l}{s_l+s_k}$ into the contribution from factor prices and the contribution from biased technical change. We believe the latter is more comparable across the two sources.

Labors share of value added in the production data declined more quickly between 1970-1999 than in the 2000s. The profit bias accounts for most of the decline before 2000. Labors share of capital/labor cost shows the opposite pattern, and is more consistent both qualitatively and quantitatively with the overall pattern using national accounts. The decline of the labor share accelerated since 2000, which is mostly accounted for by an acceleration of the bias toward capital.

D.4.4 Alternative Rental Prices

Our rental prices are based upon official NIPA deflators for equipment and structures capital. However, Gordon (1990) has argued that the NIPA deflators underestimate the actual fall in equipment prices over time. We examine how this critique might change our results on the bias of technical change by using an alternative rental price series for equipment capital that Cummins and Violante (2002) developed by extending the work of Gordon (1990). Their series extends to 1999, so we compare our baseline to these rental prices during the 1970-1999 period. Using the Cummins and Violante (2002) equipment prices implies that the wage to rental price ratio has increased by 3.8 percent per year, instead of 2.0 percent per year with

Table VIII Contributions to Labor Share Change using Production Data

Period	Labor Share Change	Contributions from: Factor Prices	Capital	Profit
	Labors Share of Value Added			
1970-1999	-0.54%	0.08%	-0.16%	-0.41%
2000-2009	-0.38%	0.05%	-0.45%	0.02%
	Labors Share of Capital and Labor Cost			
1970-1999	-0.18%	0.11%	-0.29%	
2000-2009	-0.76%	0.09%	-0.85%	

Note: All changes are annualized. The contributions from factor prices, capital/labor bias, and profit bias are as defined in the text.

the NIPA deflators. This change increases the contribution of factor prices to the labor share by about two-thirds, albeit from a low base of 0.07 percent. Given our estimate of the aggregate elasticity of substitution, changes in factor prices have not been the driving force behind the declining labor share.

Table IX Contributions to Labor Share Change with Alternative Rental Price Series

Deflator	$\frac{w}{r}$ Change	Labor Share Change	Contributions from: Factor Prices	Capital	Profit
NIPA	1.97%	-0.25%	0.07%	-0.22%	-0.10%
GCV	3.82%	-0.25%	0.12%	-0.26%	-0.11%

Note: All changes are annualized over the 1970-1999 period. The contributions from factor prices, capital/labor bias, and profit bias are as defined in the text.