

Marginal Costs in Markup Estimates*

Shuichiro Nishioka[†] and Mari Tanaka[‡]

July 24, 2023

Abstract

This paper builds an empirically tractable framework for the analysis of marginal costs in markup estimates from the production approach and examines how markups differ by the set of variable inputs. Using plant-product matched data from Japan, we show that yearly changes in markups can capture yearly changes in product prices and marginal costs, irrespective of the set of variable inputs. The production approach generally uses the most flexible intermediate inputs to compute markups. However, the estimate entails counterintuitive properties against standard models of imperfect competition because markups are computed over upward-sloping marginal cost functions. We show that the properties of markups depend crucially on how variable inputs are theoretically defined and how producers actually adjust inputs.

Keywords: Markups, Variable inputs, Marginal costs, Plant-product matched data, Japanese manufacturing

JEL Classification: D22, D24, L11

*We are grateful to Daniel Berkowitz, Kyoji Fukao, Sara E. Guffey, Dirk Krueger, and Scott Schuh for their many insightful comments. This study is conducted as a part of the project “East Asian Industrial Productivity” undertaken at the Research Institute of Economy, Trade and Industry (RIETI). This study utilizes the micro data of the questionnaire information based on “the Census of Manufacture” which is conducted by the Ministry of Economy, Trade and Industry (METI), and its converters for establishment panel and industry panel, which are provided by RIETI.

[†]John Chambers College of Business and Economics, West Virginia University, 83 Beechurst Avenue Morgantown WV 26506-0625, Tel: +1(304) 293-7875, Email: shuichiro.nishioka@mail.wvu.edu

[‡]Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan, Email: mari.tanaka@r.hit-u.ac.jp

1 Introduction

The production approach (Hall, 1988; De Loecker and Warzynski, 2012; and De Loecker et al, 2020) enables us to derive product-, plant-, or firm-level estimates of output price to marginal cost markups.¹ This methodological revolution allows us to study the impacts of the market power of individual producers on aggregate economic outcomes, including market concentration and income inequality. The approach derives markups by estimating the output elasticity of a variable input and dividing the output elasticity by the ratio of the variable input cost to revenue. Markup estimates, however, differ substantially owing to the inclusion or exclusion of a certain type of cost (Traina, 2018), returns to scale in production (Basu, 2019), and the choice of a variable input (Raval, 2023).

We argue that these differences arise because markups are computed over different marginal costs (Syverson, 2019). Marginal costs differ across producers for several reasons. Firm heterogeneity in productivity and product quality are responsible for differences in marginal costs (Hopenhayn, 1992; Kugler and Verhoogen, 2012; De Loecker et al, 2016). Even for a single producer, marginal costs could differ, depending on the choice of what is considered to be variable versus fixed inputs.

We build an empirically tractable framework for the analysis of marginal costs in markup estimates and examine the properties of markups computed using different sets of variable inputs. In this paper, the choice of variable versus fixed inputs underlies the heterogeneity in marginal costs. For example, when only the most flexible intermediate inputs are variable inputs, we implicitly assume that production functions are decreasing returns to scale in terms of variable inputs. Thus, the above-mentioned criticisms of the production approach could be associated with how variable inputs are theoretically defined to derive underlying marginal cost functions. The model predicts that marginal cost increases as input prices increase and decreases as productivity increases. Moreover, marginal cost increases (decreases) with output when the output elasticity is sufficiently less than (greater than) one.

To empirically examine whether markups from the production approach follow the theoretical predictions, we use plant-product matched data from Japan’s Census of Manufacture, an annual

¹The markup literature has grown using two distinct types of data. One strand of the literature relies on demand-side information about product prices and quantities. It computes markups from residual price elasticities of demand (e.g., Feenstra and Weinstein, 2017). The other relies on supply-side information and computes markups from the producer-side data (e.g., De Loecker and Warzynski, 2012).

survey conducted by the Ministry of Economy, Trade and Industry (METI).² There are several reasons why the data enable us to answer our research questions. First, the data contain product prices and physical output quantities.³ Without the product-level price and quantity information, we cannot examine changes or differences in markups arising from those in product prices and quantities. Second, the data cover six types of inputs: four types of intermediate inputs (materials, fuels, electricity, and outsourcing) and two types of factor inputs (labor and capital). This is ideal for us to discuss how the properties and distributions of markup estimates differ owing to the plausible adjustability of variable inputs. Lastly, Japanese producers adjust materials rapidly and labor slowly due to their practice of lifetime employment (Hashimoto and Raisian, 1985; Kambayashi and Kato, 2017). We will show that the difference in the adjustment speed of inputs is the key to explaining why markup estimates from labor and those from materials move in opposite directions in the short run (Raval, 2023).

We follow the literature (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Akerberg et al, 2015) and consider three alternative sets of variable inputs according to the plausible adjustability of inputs for computing marginal costs. First, we compute markups over marginal costs by assuming that intermediate inputs are variable inputs. This measure is frequently used in the literature (De Loecker et al., 2016; Foster et al, 2022). Second, we compute markups by assuming that intermediate inputs and labor are variable inputs without taking into account short-run labor adjustment costs (e.g., Bils, 1987). Lastly, we compute markups by assuming that all observable inputs, including capital, are variable inputs in the long run. This measure is traditionally employed in the macroeconomic literature that uses national accounts data to compute markups (e.g., Diewert and Fox, 2008).

Using product-level prices and quantities from the sample of single-product plants, we first show that yearly changes in markup estimates from the production approach can precisely capture yearly changes in product prices and marginal costs, irrespective of the set of variable inputs.⁴

²The census is conducted on all plants with more than four regular employees. The response rate is greater than 95% for each year. Thus, the census covers almost all plants in Japanese manufacturing. In this paper, we focus on the sample that covers all manufacturing plants that have 30 or more regular employees because smaller plants do not report the data on tangible capital assets and investment necessary to compute their capital stock. Although we do not include smaller plants, the data cover around 70% of Japanese manufacturing plants in terms of total shipments.

³Although we have four types of intermediate inputs (materials, fuel, electricity, and outsourcing), we do not have the prices and physical quantities information for materials. See Kugler and Verhoogen (2012) who use the Colombian manufacturing census, which contains prices and physical quantities of all inputs and outputs.

⁴We also find that cross-plant differences in markup estimates can capture those in product prices and marginal costs. Similar to De Ridder et al (2021), we find more robust evidence from changes than levels.

Conditional on the assumption that fixed inputs are truly fixed over a year, the rise in product prices and the fall in marginal costs each increase markups. Second, consistent with the theoretical predictions, marginal costs respond systematically to demand shocks, depending on returns to scale in the aggregate of variable inputs. By including only intermediate inputs in the set of variable inputs, the output elasticity tends to be sufficiently less than one, generating upward-sloping marginal cost functions.⁵ Thus, a negative demand shock for a producer tends to suppress its marginal cost and increase its markup. To examine this prediction further, we provide a case study on the Japanese semiconductor industry around the time of the dot-com bubble collapse in 2000. Intuitively, the bubble collapse should have reduced markups because of the decline in semiconductor demand; however, markups derived from intermediate inputs increased by 5.9% from 2000 to 2002. More generally, we show that upward-sloping marginal cost functions give rise to counterintuitive properties of markup estimates against standard models of imperfect competition (e.g., Atkeson and Burstein, 2008; Feenstra and Weinstein, 2017). Larger producers have lower markups, and market shares and markups are negatively associated. Our findings highlight the importance of recognizing the differences in markup estimates that arise from computing them over different marginal cost functions.

Our paper is closely related to Raval (2023) who shows that markups from materials, labor, or a composite of both are very different using data from Chile, Colombia, India, Indonesia, the United States, and Southern Europe.⁶ While Raval (2023) argues that the rise of labor-augmenting productivity in non-neutral production technology can reconcile the differences, we show that the differences arise because markups are computed over different marginal cost functions. For example, an expression of the marginal cost for labor markups (material markups) is the wage (material prices) divided by the marginal product of labor (materials). Thus, when producers can adjust only materials in the short run, the marginal products of labor and materials move in opposite directions, and so do the marginal costs and markup estimates. In particular, the discrepancy between the theoretical assumption (i.e., materials are fixed) and the empirical observation (i.e., producers adjust materials simultaneously) develops contradictory movements of labor markups.

⁵See Basu (2019) and De Loecker et al (2023) who show that scale elasticities play a crucial role in the U.S. aggregate markup trend.

⁶Traina (2018) studies the sensitivity of markup estimates due to the inclusion or exclusion of a certain type of cost. Using financial statement data from U.S. public firms, Traina (2018) shows that the U.S. aggregate markup does not increase once the marketing and management costs are included in the computation of markups.

We confirm this prediction in the case study of the semiconductor industry.

Lastly, we do not attempt to discuss how markup estimates differ due to the applications (Bond et al, 2021; De Ridder et al, 2021; and Foster et al, 2022) and methods (Akerberg et al, 2015; De Loecker et al, 2016; Gandhi et al, 2020) of estimating production function parameters.⁷ Following the popular method in the literature, we estimate industry-specific, time-invariant Cobb-Douglas production functions. We then show that markups estimated from the conventional production approach have counterintuitive properties.

The rest of the paper proceeds as follows. In the second section, we derive markups from the cost minimization problem. In the third section, we discuss the development of data. In the fourth section, we describe markup estimates. In the fifth section, we examine how the choice of variable inputs shapes marginal costs and markups. We also show how markup estimates from Japan’s semiconductor producers responded to the dot-com bubble collapse in 2000. In the last section, we discuss our conclusions.

2 Deriving Markups

In this section, we theoretically show how markups differ systematically due to the choice of variable inputs. We follow the literature and use the cost minimization problem to derive markups (Hall, 1988; De Loecker and Warzynski, 2012; and De Loecker et al, 2020). We consider the situation where a producer is a price-taker in input markets but has market power in a product market. The assumptions we impose to derive markups are (1) a producer optimizes all variable inputs but does not change fixed inputs, (2) the sum of output elasticities of variable inputs is constant over time, (3) producers do not face adjustment costs for variable inputs,⁸ and (4) the marginal product of an input is diminishing.

2.1 Production Approach

A producer i at time t uses a production function that converts inputs (q_{it}^x) into real output (Q_{it}) . The corresponding input prices (p_{it}^x) are strictly positive, exogenous for producers, and producer

⁷Markup estimates differ substantially due to revenue- or quantity-based estimates on output elasticities (Bond et al, 2021), functional forms of production functions (De Ridder et al, 2021), and the extent of industry categories (Foster et al, 2022).

⁸See, for example, Bills (1987) and Cooper and Haltiwanger (2006) who examine labor and capital adjustment costs, respectively.

specific. We use the following production function that can differ across producers and evolve over time:

$$Q_{it} = F_{it}(\cdot) = \Omega_{it} \prod_{x \in V} (q_{it}^x)^{\alpha_{it}^x} \prod_{x \in K} (q_{it}^x)^{\alpha_{it}^x} \quad (1)$$

where the output elasticity of an input x is

$$\alpha_{it}^x = \frac{\partial Q_{it} / Q_{it}}{\partial q_{it}^x / q_{it}^x}. \quad (2)$$

Here, an input x is either in V , a set of variable inputs, or in K , a set of fixed inputs. And, Ω_{it} is a Hicks-neutral productivity measure that is a source of producer-level heterogeneity in marginal cost.⁹

A producer can freely adjust quantities of variable inputs (q_{it}^x where $x \in V$) at any point in time without incurring any adjustment costs but cannot change quantities of fixed inputs (q_{it}^x where $x \in K$). Therefore, in the following Lagrangian function, we assume that a producer is able to optimize variable inputs only ($x \in V$):

$$\mathcal{L} = \sum_{x \in V} p_{it}^x q_{it}^x + \sum_{x \in K} p_{it}^x q_{it}^x + \lambda_{it}^V [Q_{it} - F_{it}(\cdot)]$$

where the Lagrangian multiplier (λ_{it}^V) is the marginal cost to produce an exogenous quantity of real output (Q_{it}).

The assumption we impose on the cost minimization problem above is that (1) producers optimize all variable inputs; and (2) the sum of output elasticities of variable inputs is constant over time:

$$\sum_{x \in V} \alpha_{it}^x = \alpha_i^V$$

where the output elasticity (α_i^V) increases as the scope of variable inputs widens (V) by treating more inputs as variable inputs according to the adjustability of inputs the literature generally assumes (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Akerberg et al, 2015).

⁹See Doraszelski and Jaumandreu (2018) and Raval (2023) who allow labor-augmenting productivity. Our theoretical predictions are general to labor-augmenting productivity.

The first-order condition of a variable input is

$$\lambda_{it}^V = \frac{p_{it}^x q_{it}^x}{\alpha_{it}^x Q_{it}}. \quad (3)$$

Using first-order conditions of all variable inputs, we can find the optimal quantities of variable inputs to produce Q_{it} . Then, we can derive the following formula for marginal cost:

$$\lambda_{it}^V = \frac{\sum_{x \in V} p_{it}^x q_{it}^x}{\alpha_i^V Q_{it}}. \quad (4)$$

Finally, we develop markups from the product price (P_{it}) divided by marginal cost (λ_{it}^V):

$$\mu_{it}^V = \alpha_i^V \frac{P_{it} Q_{it}}{\sum_{x \in V} p_{it}^x q_{it}^x}. \quad (5)$$

Equation (5) is the production approach of estimating markups from the output elasticity of the aggregate of variable inputs divided by the ratio of variable input cost to revenue.

2.2 Variable Inputs and Marginal Costs

To understand how marginal costs in markup estimates are associated with the set of variable inputs, we prepare an alternative expression of marginal cost (e.g., Roeger, 1995):

$$\lambda_{it}^V \approx (\Omega_{it})^{-1/\alpha_i^V} (Q_{it})^{1/\alpha_i^V - 1} p_{it}^V q_{it}^K \quad (6)$$

where $p_{it}^V = \prod_{x \in V} (p_{it}^x)^{\alpha_{it}^x / \alpha_i^V}$ is the weighted average of variable input prices, and $q_{it}^K = \prod_{x \in K} (q_{it}^x)^{-\alpha_{it}^x / \alpha_i^V}$ is the composite of fixed inputs, which is assumed to be constant over a year.¹⁰

Equation (6) is a dual form of equation (1), which shows that marginal cost differs across producers with productivity, real output, variable input prices, and quantities of fixed inputs. Marginal cost increases as input prices increase and declines as productivity increases. Conditional on other variables, marginal cost also varies systematically with output when $\alpha_i^V \neq 1$. Consider the following two polar cases: $\alpha_i^V \ll 1$ when the most flexible inputs (i.e., materials) are the only variable inputs in the short run, and a production function is decreasing returns to scale in variable inputs; and $\alpha_i^V \gg 1$ when all the inputs are variable inputs in the long run, and a production

¹⁰We suppress a constant term in equation (6).

function is increasing returns to scale. In the former case, we have an upward-sloping marginal cost function with respect to output because $1/\alpha_i^Y - 1$ in equation (6) is positive. As a result, conditional on other variables, larger producers have higher marginal costs than smaller producers, and positive demand shocks increase marginal costs. In the latter case, we have a downward-sloping marginal cost function. Thus, larger producers have lower marginal costs than smaller producers, and positive demand shocks decrease marginal costs.

Raval (2023) shows that markups from materials, labor, or a composite of both are very different. While Raval (2023) finds that non-neutral technology can reconcile the differences,¹¹ we examine if the differences arise because markups are computed over different marginal cost functions. As a simple example, consider a two-input (materials (M_{it}) and labor (L_{it})) constant returns to scale Cobb-Douglas production function ($\alpha_i^M = \alpha_i^L = 1/2$)¹² with exogenous input prices (material price (p_{it}) and wage (w_{it})). We follow Raval (2023) and consider markups from materials, labor, or a composite of both. First, we derive the marginal cost function from the first-order condition of materials by fixing labor:

$$\lambda_{it}^1 \approx (\Omega_{it})^{-2} Q_{it} p_{it} L_{it}^{-1}. \quad (7)$$

Second, we derive the marginal cost function from the first-order condition of labor by fixing materials:

$$\lambda_{it}^2 \approx (\Omega_{it})^{-2} Q_{it} w_{it} M_{it}^{-1}. \quad (8)$$

Lastly, we use the first-order conditions of materials and labor to compute markup. Then, the marginal cost function is the weighted input prices divided by productivity:

$$\lambda_{it}^3 \approx (\Omega_{it})^{-1} w_{it}^{0.5} p_{it}^{0.5}. \quad (9)$$

The marginal cost functions from equations (7), (8), and (9) are very different. When markups are derived from materials in equation (7) or labor in equation (8), the marginal cost functions are sensitive to changes in the targeted level of output. Moreover, these estimates depend crucially on how producers adjust inputs when they experience demand shocks. While equation (8) is derived under the assumption that materials are fixed inputs, producers actually adjust materials

¹¹When labor and materials are complements, higher labor-augmenting productivity would lower labor's output elasticity relative to materials' output elasticity.

¹²The parameter setting is an illustrative example, and a similar conjecture holds in general.

rapidly and labor slowly. The discrepancy between the theoretical assumption and the empirical observation explains why λ_{it}^1 and λ_{it}^2 move in opposite directions. To better understand why, consider the following expressions of marginal costs from equation (3):

$$\lambda_{it}^1 = \frac{p_{it}}{MP_{it}^M} \quad \text{and} \quad \lambda_{it}^2 = \frac{w_{it}}{MP_{it}^L}$$

where the marginal products of materials and labor are

$$MP_{it}^M = 0.5\Omega_{it} \left(\frac{L_{it}}{M_{it}} \right)^{0.5} \quad \text{and} \quad MP_{it}^L = 0.5\Omega_{it} \left(\frac{M_{it}}{L_{it}} \right)^{0.5}.$$

The marginal cost functions above indicate that λ_{it}^1 and λ_{it}^2 move in opposite directions through the inverse association between materials and labor in marginal products. Our findings later in the paper confirm that facing the collapse of the dot-com bubble in 2000, Japan's semiconductor producers reduced materials rapidly and labor slowly. While markups derived from materials increased by 6.5% from 2000 to 2002, markups derived from labor declined by 11% in the same period.

Although the assumption that producers are unable to adjust labor is plausible in the short run, the upward-sloping marginal cost function in equation (7) gives rise to counterintuitive properties of markup estimates. Consider a case where a producer faces a negative demand shock, and its product price and the targeted level of output decline simultaneously. In this case, keeping labor constant, the producer can only decrease materials to cut down production. As a result, the marginal product of materials increases, and the marginal cost declines. Markups could increase or decrease depending on the relative magnitude of the fall of product price versus the fall of marginal cost. Consider another case where a producer gains market power and raises its price by scaling down output. In this case, the producer's demand for materials declines, its marginal product of materials increases, and its marginal cost declines. Therefore, the markup increases more than the product price increases.

3 Data

3.1 The Census of Manufacture in Japan

The Census of Manufacture is an annual survey conducted by the Ministry of Economy, Trade and Industry (METI). The data contain the product prices and physical quantities at the product level (e.g., Foster et al, 2008; Kugler and Verhoogen, 2012; De Loecker et al, 2016). Without the data on product-level price and quantity information, we cannot examine how yearly changes in markups are associated with yearly changes in product prices and marginal costs. Importantly for the purpose of this paper to study the adjustability of variable inputs, the data cover four types of intermediate inputs (materials, fuel, electricity, and outsourcing) and two types of factor inputs (labor and capital). Thus, we can examine the combinations of variable versus fixed inputs.

The data consist of two layers: plant-level and product-level. Plant-level variables, such as revenues, employment, wage bills, spending on intermediate inputs, and investments, are from the plant-level data set; and product-level variables, such as shipments and physical quantities (which give information on unit prices), are from the product-level data set. In the analysis, we use only the sample that covers all manufacturing plants that have 30 or more employees. We exclude plants with less than 30 employees because these plants are not required to report the data on tangible capital assets and investment necessary to compute their capital stock.¹³ We also drop the top and bottom 1% of observations for each markup estimate for each year as outliers.

Over the period 1987-2009, we have 987,299 observations at the plant level. On average, we have around 43,000 plants for each year. The sample size of product-level data that includes information on prices and quantities, however, is smaller. The information on product is reported in METI's six-digit product classification system. There are approximately 2,000 products, of which quantity information is available for around 800 products. After we merge the product-level data with the plant-level data, we have 264,740 plant-year observations. The number of plant-year observations decreases further to 67,125 once we limit the sample to single-product manufacturing plants.¹⁴

Table 1 reports the summary statistics of the plant-level data for the years 1987 and 2007. The

¹³We do not include the resource-intensive industries (tobacco, oil, and coal refinery) and focus on the remaining 49 industries (the industry category defined by the Japan Industry Productivity (JIP) 2015 database).

¹⁴The sample size declined substantially because 48% of the sample were single-product establishments, and 14% reported product quantities. Despite this, the sample of single-product plants almost proportionally covered a wide range of industries (see Fukao and Ito, 2010).

data reported in this table do not include product-level information.¹⁵ One notable finding in the table is that while the mean values of log real revenues, log real spending on intermediate inputs, and log real capital stock increased substantially, the mean value of log labor did not change over the period. On average, real revenues, real spending on intermediate inputs, and real stock of capital increased by 44.4%, 31.1%, and 66.9%, respectively.¹⁶ The average size of employment, however, increased only minimally by 5.5%. Our findings may reflect Japanese producers' practice of lifetime employment (Hashimoto and Raisian, 1985). Even during the downturn in the 1990s after the collapse of the real estate bubble economy, Japanese firms were unable to adjust labor rapidly to changing market conditions (Kambayashi and Kato, 2017).

3.2 Cost Shares

This section uses cost shares to examine the input structure in Japanese manufacturing. The data contain payments for four types of intermediate inputs (materials, fuel, electricity, and outsourcing)¹⁷ and labor.¹⁸ Because the data do not report capital stock and capital cost, we develop capital stock from the perpetual inventory method and estimate capital cost from the opportunity cost of holding capital as assets (see Appendix I). This *ex-ante* approach to compute capital cost was proposed by Jorgenson and Griliches (1967) and has been applied to studies including Caballero and Lyons (1992), Barkai (2020), and De Loecker et al (2020).

The cost share of an input x in total costs is

$$cs_{it}^x = \frac{p_{it}^x q_{it}^x}{\sum_{x \in V, K} p_{it}^x q_{it}^x}.$$

¹⁵See Table A3 in the Appendix for the summary statistics of the product-level data.

¹⁶There are several potential reasons why the capital-labor ratio surged during the period. First, Japanese manufacturing firms over-invested in tangible, reproducible capital during the bubble economy in the late 1980s and the early 1990s. Second, the Japanese government's bank bailouts in 1998 and 1999 injected capital into the real economy (Giannetti and Simonov, 2013). Indeed, while the nominal wage increased by 25%, the capital service price declined by 28% from 1987 to 2007.

¹⁷The data include payments for (1) materials, (2) fuel, (3) electricity and (4) outsourcing. First, the cost of raw materials represents the total consumption of major raw materials, auxiliary supplies, purchased components and parts, containers, packing materials, and plant maintenance materials. Second, the cost of fuel includes coal and petroleum expenses, including private power generation. Third, the cost of electricity represents total payments for power supply by vendors, excluding private power generation. Lastly, outsourcing represents various payments for outsourcing and subcontracting. This category includes payments made and accounts payable to subcontractors for consigned production and processing and payments regarding services such as repair, inspection and maintenance of production equipment, operation of machinery.

¹⁸The value of total wages and salaries is defined as the total amount of basic wages, basic allowances, special allowances (e.g., year-end bonus) paid to employees among regular and part-time workers, and other allowances. Other allowances include retirement allowances and discharge allowances for employees, payments to workers dispatched from other companies, wages for temporary workers, and payments to workers dispatched to other companies.

Cost shares have important implications for production technologies because cost shares could approximate output elasticities at the producer level when production functions are constant returns to scale and producers optimize all inputs. For example, if a producer’s production function is an industry-level Cobb-Douglas form, then cost shares and output elasticities are common across producers in an industry. However, if the production function deviates from the form, cost shares could be different across producers, depending on various factors, including computer investment and diversification to non-manufacturing activities (Foster et al, 2022).

Table 2 reports the summary statistics of cost shares for the years 1987 and 2007. The mean value of cost shares of labor declined by 1.5 percentage points over the period from 30.3% in 1987 to 28.8% in 2007. The average cost share of capital increased slightly from 5.3% in 1987 to 5.7% in 2007, and those of intermediate inputs increased from 64.4% in 1987 to 65.5% in 2007. The data in Table 2 suggest that Japanese manufacturing used slightly less labor and more capital and intermediate inputs over the period.

Note that there is a large variation in cost shares across manufacturing plants. For example, the standard deviation of cost shares of materials was 24.5% in 1987 and 23.5% in 2007. The large variation in cost shares generally suggests that producers use different production technologies. The large variation in cost shares of capital also implies that adjustment costs in capital prevent the efficient allocation of capital across producers (Hsieh and Klenow, 2009; Asker et al, 2013).

4 Markup Estimates

4.1 Empirical Strategy

We use equation (5) and prepare markups using the three sets of variable inputs. First, we assume intermediate inputs are variable inputs to derive markups. This measure is frequently used because the literature recommends using a first-order condition of the most flexible input to derive markups.¹⁹ Second, we derive markups over marginal costs using intermediate inputs and labor.²⁰ Finally, we assume that all observable inputs, including capital, are variable inputs.

¹⁹See Assumption 2 in De Loecker et al (2016, p455). They assume that the production function is continuous and twice differentiable with respect to at least one static (i.e., freely adjustable or variable) input. De Loecker and Warzynski (2012) emphasize that using the conditional cost function without fully optimizing all of the inputs can prevent consideration of the full dynamic problem and impose additional assumptions to derive markups.

²⁰De Loecker et al (2020) prepare a specification that treats labor as a variable input, and De Loecker and Warzynski (2012) use the first-order condition of labor to derive markups.

The identification and estimation of production function parameters are challenging for several reasons. First, Bond et al (2021) show that output elasticity estimates using revenue-based output data differ greatly from those using quantity-based output data. Second, Foster et al (2022) find that output elasticities of materials decline with more detailed categories of industries and are smaller for larger firms with advanced technologies.

Our primary objective is to understand the properties of markups estimated from the conventional production approach. As such, we follow the literature initiated by De Loecker and Warzynski (2012) and estimate industry-level output elasticities to compute markups. In this case, we use the output and input deflators and add controls for the joint output and input price term to estimate output elasticities.²¹ After estimating output elasticities, we approximate the output elasticities of the aggregate of variable inputs: $\alpha_i^V = \alpha^M$ when variable inputs are intermediate inputs (i.e., materials, fuel, electricity, and outsourcing); $\alpha_i^V = \alpha^M + \alpha^L$ when variable inputs are intermediate inputs and labor; and $\alpha_i^V = \alpha^M + \alpha^L + \alpha^K$ when all inputs (intermediate inputs, labor, and capital) are variable inputs. The industry-level output elasticities of intermediate inputs, labor, and capital are on average 0.59, 0.36, and 0.08, respectively (see Tabel A2 in the Appendix). Thus, α_i^V increases from 0.59 to 0.95, and 1.04 as the scope of variable inputs widens according to the adjustability of inputs.

The assumption that all plants in an industry share the same production technology is inconsistent with the large variation in cost shares in Table 2. To avoid spurious results from potentially biased estimates and presumably imprecise assumptions on the output elasticities, we use first-differenced log markups so that the unobserved plant-specific, time-invariant output elasticity (α_i^V) in equation (5) would be irrelevant:

$$\Delta \ln(\mu_{it}^V) = \ln(\mu_{it}^V) - \ln(\mu_{i,t-1}^V). \quad (10)$$

De Loecker (2020) and De Ridder et al (2021) show that measurement errors in output elasticities affect the level of markups but do not severely affect the yearly change in log markups.

²¹See Appendix II.

4.2 Markup Trend in Japan’s Manufacturing

Several papers examine markup trends in Japan over the period we study. While Kiyota et al (2009) find a declining trend in markups, our results suggest that markups were stable or slightly increasing over the period. Karabarbounis and Neiman (2018) and De Loecker and Eeckhout (2020) show increasing markup trends in Japan from public firms as a part of their studies on global markup trends.

Figure 1 illustrates the mean and median values of markups over the period of 1987-2009. Here, we use markups when intermediate inputs and labor are variable inputs by implicitly assuming that labor is adjustable in the long run (see the corresponding summary statistics in columns (3) and (4) in Table 3). The figure shows that markups declined slightly over the burst of the real estate bubble economy in the early 1990s, then increased in the early 2000s, and declined over the mid-2000s. Overall, markups were stable in Japan: the mean value of markups was around 1.4, and the median value was around 1.3 throughout the period.²² Thus, we do not find a sharp rise in markups, as in the United States (De Loecker et al, 2020; Autor et al, 2020), in Japanese manufacturing.²³

4.3 Variable Inputs and Markup Estimates

Table 3 reports the summary statistics of the levels of markup estimates for the years 1987 and 2007. Columns (1) and (2) report markups computed over marginal costs when intermediate inputs are variable inputs, columns (3) and (4) report markups when intermediate inputs and labor are variable inputs, and columns (5) and (6) report markups when we treat all observable inputs as variable inputs. The first row reports the weighted means (i.e., weighted by revenues), and the remaining rows report the summary statistics of the distributions.

We first discuss the production size implications for markup estimates. When we assume that intermediate inputs are the only variable inputs and the output elasticities of intermediate inputs are around 0.59, then industry-specific marginal cost functions would be upward-sloping, and larger plants in an industry could have higher marginal costs and lower markups.²⁴ If this is the case,

²²When we regress log product price with log physical output quantity, we find statistically significant, negative correlations across different specifications. Thus, our finding that markups are greater than one for most of the plants is consistent with the demand-side information that plants do not face perfectly elastic residual demand curves.

²³The trends of markups, however, are very different across industries. We report the results from the Melitz and Polanec (2015) decomposition in Appendix III for the entire manufacturing sector and several selected industries.

²⁴The marginal cost function from equation (6) is positively associated with real output and negatively associated

the weighted mean could be lower than the unweighted mean. Table 3 supports this argument only if intermediate inputs are treated as variable inputs (the weighted mean is 1.095, and the unweighted mean is 1.354 in 1987, as in column (1)). When we add labor as variable inputs, the output elasticities of variable inputs are close to one. Then, as is theoretically shown in equation (6), marginal cost becomes less sensitive to real output. Consistent with this intuition, the weighted means are slightly greater than the unweighted means (the weighted mean is 1.436, and the unweighted mean is 1.369 in 1987, as in columns (3)). When we further add capital as a variable input, the results do not change primarily because the contribution of capital stock in output is small.

Figure 2 illustrates the unconditional correlation between markups and log real output in 1997. We use binned scatter plots and fitted lines to visualize relationships between log real revenue and markups. Consistent with the theoretical discussions in Section 2 and summary statistics in Section 4, while there is a negative correlation when we treat intermediate inputs as variable inputs, we do not find such a strong negative correlation when we add labor and then capital as variable inputs.

Raval (2023) illustrates that markups from materials are more dispersed than markups from a composite of both. Consistent with Raval’s (2023) findings, Table 3 shows that standard deviations are the largest for markups over marginal costs using intermediate inputs (0.921 in 1987), and decline as we include labor (0.378 in 1987) and then capital (0.370 in 1987). Consistent with equation (6), as α_i^V declines from unity, the variance of markup estimates increases because that of marginal costs conditionally increases with real output.

5 The Properties of Markup Estimates

5.1 Price, Marginal Cost, and Markup Dynamics

The results in Table 3 and Figure 2 illustrate how the choice of variable versus fixed inputs could change markup estimates through underlying marginal cost functions. The results, however, could depend on the assumption that all plants in an industry share the same production technology. To avoid spurious results from potentially biased estimates of output elasticities, we next examine the relationship implied in equations (5) and (6) that relates yearly changes in a plant’s markup to

with fixed inputs. Thus, it is not entirely clear how marginal costs differ across producers due to their size because large producers tend to use more fixed inputs.

yearly changes in product price, output quantity, productivity, and variable input price:

$$\Delta \ln(\mu_{it}^V) = \Delta \ln(P_{it}) + (1 - 1/\alpha_i^V) \Delta \ln(Q_{it}) + (1/\alpha_i^V) \Delta \ln(\Omega_{it}) - \Delta \ln(p_{it}^V). \quad (11)$$

There are several reasons why we use log differences and examine yearly changes in markups. First, we use plant-product matched data and apply product prices and quantities in equation (11). Because prices and quantities are different across plants even within a narrowly defined product group due, for example, to product quality (Kugler and Verhoogen, 2012), we use first differences and focus on changes in prices and quantities. Second, we can only estimate output elasticities at the industry level. By using yearly changes, we can difference out not only plant-specific time-invariant components (i.e., the quantities of fixed inputs) but also potentially biased output elasticity estimates.

We estimate equation (11) for markups estimated from the three sets of variable inputs. As for the measure of prices and quantities, we have two alternatives. One is to use the product prices and quantities in product-level data by focusing on plants that produce a single product. The other is to use the industry deflators. Measures of productivity depend on estimated industry-level output elasticities and measures of real output. When we use physical quantities, we use a quantity-based productivity measure. When we use revenues and deflators, we use a revenue-based productivity measure. Finally, variable input price is the weighted average of variable input prices. We use the plant-level wage²⁵ and the industry-level deflators with the weights from estimated output elasticities and cost shares. For example, when we derive the variable input price from intermediate inputs and labor, it is the weighted average of the industry-level input deflator for intermediate inputs (the weighted average of prices for materials, fuel, electricity, and outsourcing) and plant-level wage. Table A3 in the Appendix reports the summary statistics of the variables we use to estimate equation (11).

Table 4 reports the results. Columns (1) to (3) report the results when we use product-level prices and quantities from single-product plants, and columns (4) to (6) report the results when we use industry-level deflators. Over the period 1987-2009, we have 57,325 first-differenced observations from the single-product plants and 870,935 first-differenced observations of the entire sample of plants. We use the ordinary least squared (OLS) specification using the first differences for all

²⁵The plant-level wage is normalized by the base-year (2000) aggregate value of wage.

variables. In all of the equations, the coefficients of the log of price are positive, statistically significant, and close to one for both the product- and industry-level measures of prices. The coefficients on log variable input price are negative, statistically significant, and close to negative one for both the product- and industry-level measures of prices.²⁶ The results indicate that changes in markup estimates from the production approach precisely capture changes in output and variable input prices. The coefficients on output and variable input prices are slightly weaker in magnitude, around less than 0.8 when we use industry-level deflators.

According to equation (11), the coefficients on changes in log real output should reflect the output elasticities in the aggregate of variable inputs: $1 - 1/\alpha_i^V$. On average, α_i^V is around 0.59 when intermediate inputs are the only variable inputs, α_i^V is around 0.95 when intermediate inputs and labor are variable inputs, and α_i^V is around 1.04 when all inputs are variable inputs. From column (1) to column (3) in Table 4, the estimated coefficient on changes in physical quantities is -0.45, -0.07, and -0.014, respectively. These are close to the theoretically predicted average coefficients from $1 - 1/\alpha_i^V$: -0.67, -0.05, and 0.04, respectively.²⁷ The coefficient on real output is a large, negative value when the output elasticity is sufficiently less than one, and the production function is decreasing returns to scale in variable inputs. Furthermore, it becomes less sensitive to the change in real output when the output elasticity is closer to one by including labor, then capital. The coefficient on changes in log productivity should also reflect the output elasticity: $1/\alpha_i^V$. Consistently, the estimated coefficient declines as the scope of variable inputs widens. Overall, markups derived from the production approach are consistent with the theoretical predictions. The estimates, however, are sensitive to the handling of variable versus fixed inputs because returns to scale in variable inputs and underlying marginal cost functions differ according to the scope of variable inputs.

Table 5 reports the results when we estimate equation (11) in levels with industry fixed effects. The results are consistent with those reported in Table 4 with relatively smaller coefficients on all the variables. We suspect that there are several reasons why estimated coefficients are smaller with log markups. In particular, the implicit assumption that fixed inputs are fixed over the period could be responsible for this tendency because producers adjust labor and capital in the long run.

²⁶Our estimates are less than one on output and input prices across all specifications. This could be attenuation bias due to measurement errors.

²⁷For example, $1 - 1/\alpha_i^V = 1 - 1/0.59 = -0.67$.

5.2 Demand Shock Analysis

In the previous section, we provided correlational evidence that, conditional on the assumption that fixed inputs are truly fixed, markup estimates from the production approach are positively associated with product prices and negatively associated with marginal costs. To show that markup estimates with a different set of variable inputs respond differently to exogenous shocks, we examine the Japanese semiconductor industry over the dot-com bubble collapse in 2000.²⁸ The information technology and telecommunications industries grew rapidly in the 1990s due to the massive adoption of personal computers and the Internet. Between 1995 and 2000, the Nasdaq index rose by more than fivefold; however, by 2002, it declined substantially, erasing all its gains during the bubble.

The dot-com bubble collapse in the United States was an unexpected, negative shock to Japan's semiconductor industry. For example, Gartner Group reported on October 9, 2000, that worldwide semiconductor sales would show double-digit growth in the next three years as manufacturers found places for them in a variety of devices other than personal computers (The Associated Press). Despite the optimism from industry analysts, the dot-com bubble collapse substantially impacted the semiconductor industry. Figure 3 uses the balanced sample of 591 establishments in the semiconductor industry over the period from 1998 to 2003 and illustrates that, on average, real semiconductor sales declined by 27% from 2000 to 2002. Consistent with the discussions in Akerberg et al (2015), Figure 3 illustrates that the Japanese manufacturers adjusted materials rapidly and labor slowly. Table 6 reports the changes in all inputs from 2000 to 2002. The spending on materials declined significantly by 35%, whereas labor declined by 15% and capital stock declined by 17%. And, somewhat surprisingly, spending on fuels and electricity declined minimally by 12% and 10%, respectively.²⁹ Semiconductor manufacturers responded to the negative shock by reducing all inputs.³⁰

Intuitively, the dot-com bubble collapse should have reduced markups because of the decline in semiconductor demand. Indeed, the average price of semiconductors declined over the period.³¹

²⁸See Syverson (2004) and Collard-Wexler (2013) who use a case study of the U.S. concrete industry to study the empirical association between demand and competition. The semiconductor industry is a representative case of Japanese manufacturing in terms of output elasticities. The output elasticities of intermediate inputs, labor, and capital are 0.53, 0.39, and 0.08, respectively.

²⁹This is probably because semiconductor producers are required to maintain, for example, energy-intensive clean-rooms regardless of output volume.

³⁰To show that the dot-com bubble collapse did not impact other manufacturing industries, we report the results from the automobile industry in Figures A1 and A2 in the Appendix.

³¹The semiconductor prices have declined constantly over a long period of time primarily because of advancements

Figure 4, however, illustrates that markups derived from intermediate inputs increased by 5.9% from 2000 to 2002. The findings from the semiconductor industry are consistent with our theoretical predictions. When we use a narrow scope of variable inputs, and the output elasticity of variable inputs is substantially less than one, a negative demand shock leads to an increase in the marginal product of materials, which reduces marginal costs and increases markups. We do not find a similar trend when we use intermediate inputs and labor as variable inputs: on average, the markup estimates decreased by 1.1% from 2000 to 2002. Although the literature recommends computing markups from the most flexible intermediate inputs, the cyclical movement of marginal costs makes the interpretation of the markup trend difficult when we follow the production approach and compute markups from the most flexible intermediate inputs.

We theoretically showed that, when producers adjust materials rapidly and labor slowly, the marginal products of labor and materials move in opposite directions; hence, so do labor markups and material markups. Figure 4 confirms that markup estimates from labor declined substantially by 11% from 2000 to 2002. We find similar trends in markups derived from sub-categories of intermediate inputs.³² When we use the most inflexible energy (fuels and electricity) to derive markups, energy markups declined by 17.1%. When we use the most flexible materials to derive markups, material markups increased by 6.5%. The movements of markups depend crucially on how variable inputs are theoretically defined to derive markups.

To highlight the movement of marginal cost from labor and that of materials, Figure 5 reports the log difference between markups from intermediate inputs and labor and those from intermediate inputs,

$$\Delta \ln(\mu_{it}^{ML}) - \Delta \ln(\mu_{it}^M) = \ln(\lambda_{it}^M / \lambda_{it}^{ML}), \quad (12)$$

and, the log difference between markups from intermediate inputs and labor and those from labor,

$$\Delta \ln(\mu_{it}^{ML}) - \Delta \ln(\mu_{it}^L) = \ln(\lambda_{it}^L / \lambda_{it}^{ML}). \quad (13)$$

By making the log difference between two measures of markups, we can difference out plant-
in production technologies.

³²The output elasticities are computed from the estimated output elasticity of intermediate inputs and cost shares within intermediate inputs.

specific product prices and focus on the relative changes in marginal costs. Figure 5 illustrates that relative markups from equations (12) and (13) move in opposite directions. The positive demand shock in 2000 increased marginal costs from intermediate inputs relative to those from intermediate inputs and labor, and the negative demand shock in 2001 and 2002 decreased marginal costs from intermediate inputs relative to those from intermediate inputs and labor. Importantly, the positive demand shock in 2000 decreased marginal costs from labor relative to those from intermediate inputs and labor, and the negative demand shock in 2001 and 2002 decreased marginal costs from labor relative to those from intermediate inputs and labor. Thus, while markup estimates from the production approach are consistent with theoretically derived marginal cost functions, changes in marginal costs depend crucially on how producers actually adjust inputs.

5.3 Markups and Market Shares

The standard imperfect competition models generally predict that a producer’s market share and its markup are positively associated. This theoretical linkage is important to understand because, for instance, the literature emphasizes that increases in superstars’ market shares are key to explaining the rise of aggregate markups in the United States (De Loecker et al, 2020; Autor et al, 2020). As such, this section examines the empirical associations between markup estimates and market shares.

Table 7 reports the results. Columns (1) through (3) report the results when we define each of industries as a market, and columns (4) through (6) report the results when we define each of industry-prefecture pairs as a market.³³ If the market competition models mentioned above shape market shares, we should expect that market shares and markups are positively associated. The results reported in columns (1) and (4), however, show that markups and market shares are negatively associated when intermediate inputs are variable inputs. The results reported in columns (2) and (5) indicate that markups and market shares are positively associated when intermediate inputs and labor are variable inputs. We also find positive correlations when capital is additionally included in the set of variable inputs. Not surprisingly, the results imply that aggregation could be sensitive to how variable inputs are selected and how marginal costs are derived because marginal

³³Admittedly, our definition of a market (the industry or industry $\times$ prefecture segment) is not perfect because the market definition is fixed over time. Nonetheless, we can show the systematic results according to the handling of variable inputs.

costs are positively associated with real output when the output elasticities of the aggregate of variable inputs are substantially less than one.

6 Conclusion

The production approach enables us to derive output price to marginal cost markups at the product, plant, or firm level. This methodological revolution allows us to study the impacts of the rising market power of individual producers on aggregate outcomes. Using plant-product matched data from Japanese manufacturing surveys, we examined how markups differ by the adjustability of variable inputs and showed that yearly changes in markups from the production approach precisely capture yearly changes in product prices and marginal costs. The properties and distributions of markups, however, are sensitive to how variable and fixed inputs are theoretically defined to derive underlying marginal cost functions. In particular, marginal costs in markup estimates are positively associated with output when only the most flexible intermediate inputs are selected as variable inputs, which gives rise to counterintuitive properties of markup estimates against standard models of imperfect competition. Applied researchers should recognize markups computed over different marginal costs, examine how producers actually adjust variable and fixed inputs, and establish robustness using relevant sets of variable inputs to compute markups.

Appendix

I. Capital Costs

Consider that K_{it} is the quantity of capital stock that plant i owns in year t , and r_t is its corresponding price for the capital stock. We also introduce the following notation: I_{it} is the quantity of the investment good newly acquired to produce good i in time t , and p_t is the corresponding price of the investment good. Following the perpetual inventory method, the accumulated stock of past investments has the following property:

$$K_{it} = (1 - \delta)K_{i,t-1} + I_{it} \quad (14)$$

where δ is the depreciation rate of the investment good.

We have the accounting value of tangible assets from (1) non-residential buildings and structures; (2) machinery and equipment; and (3) transport equipment at the plant level. We also have investment values from these three distinct tangible assets. To obtain capital stock at the plant level, we first use the perpetual inventory method and develop total capital stock from the real investment values. We then allocate total capital stock to each plant by its share of the accounting value of tangible assets. This value is used only for the year when the plant first appears in the data. Second, we take the plant-level first-year value of capital stock and real investment values. We then apply equation (14) to obtain capital stock after the first year.

Total capital stock value at any point in time is around 40-45% of the corresponding capital stock in the Japan Industry Productivity (JIP) database 2015. There are several reasons why our capital stock is smaller. First, our sample is limited to include only plants with more than 30 employees, and we drop the three industries (tobacco, oil refinery, and coal refinery). Second, the JIP database covers a broader set of assets including some intangible assets. Although our estimates do not perfectly match the JIP database, we have an aggregate trend that is similar to the JIP database.

Next, we follow Jorgenson and Griliches (1967) and compute capital costs from the opportunity costs of holding the capital stock:

$$r_t K_{it} = p_t K_{it} (i_t - \Delta p_t / p_t + \delta) \quad (15)$$

where i_t is the risk-free interest rate (i.e., the government bond rate derived from the International Financial Statistics of the International Monetary Fund), and $\Delta p_t/p_t$ is the rate of capital gain or loss on the capital stock.

Equation (15) should capture the opportunity cost of holding versus investing capital stock as an asset. If a plant does not use its capital stock, the plant can sell it at the current market price, invest it in a risk-free asset, and collect interest payments, but lose an opportunity to gain from the potential appreciation of the capital asset. In addition, the producer can avoid losing capital asset value from physical depreciation. Because the data do not contain detailed data on debts and borrowings at the plant level, we are not able to develop Hall and Jorgenson’s (1967) measure of capital costs that take account of debt and equity financing and business income tax.

II. Estimating Output Elasticities

To obtain output elasticities, we estimate production functions at the industry level. Here, we assume the Cobb-Douglas production function that does not restrict returns to scale:

$$Q_{it} = \Omega_{it} (M_{it})^{\alpha^M} (L_{it})^{\alpha^L} (K_{it})^{\alpha^K} \quad (16)$$

where we aggregate inputs into the three categories: intermediate inputs (M_{it}),³⁴ labor (L_{it}), and capital (K_{it}).

We follow an approach proposed by De Loecker et al (2016) and obtain the output elasticities of inputs and the unobserved input price bias parameter for the 49 manufacturing industries. De Loecker et al (2016) control for input price variations across plants using the information on plant-level output prices because producers of more expensive products use more expensive inputs (Kugler and Verhoogen, 2012). Since our sample declines substantially when we use the direct measure of output prices, we follow their intuition and approximate unobserved input price biases by market shares.

To estimate the production function at the industry level, we use the timing assumption in Akerberg et al (2015) that firms need more time to optimize labor and install capital than to purchase intermediate inputs. It follows from this timing assumption that a plant’s demand for

³⁴The real spending on intermediate inputs is constructed from the four types of deflators from the JIP database (Fukao et al, 2007). For example, spending on electricity is deflated by output price deflator of electricity production.

intermediate inputs depends on its productivity and the predetermined amounts of labor and the current stock of capital.³⁵ We also follow De Loecker et al (2016) and handle unobserved input price biases with log domestic market share (s_{it}):

$$m_{it} = h_t(\omega_{it}, l_{it}, k_{it}, s_{it})$$

where lower-case variables represent the logged values (e.g., $l_{it} = \ln(L_{it})$).

Following Akerberg et al (2015), we assume the equation above can be inverted with productivity:

$$\omega_{it} = h_t^{-1}(m_{it}, l_{it}, k_{it}, s_{it}).$$

We then approximate q_{it} with the second-order polynomial function of the three inputs and interact it with the variable for input price biases:

$$q_{it} \approx \Phi_t(m_{it}, l_{it}, k_{it}, s_{it}) + \epsilon_{it}. \quad (17)$$

Next, we obtain the predicted value of equation (17), $\hat{\Phi}_t$, and compute the corresponding value of productivity for any combination of parameters Ω . We need to estimate not only a constant term and the output elasticities of the three inputs (α^M , α^L and α^K), but also the unobserved input price bias parameter, the interaction of the market share s_{it} with m_{it} (β). This enables us to express the log of productivity as follows:

$$\bar{\omega}_{it}(\Omega) = \hat{\Phi}_t - (c_j + \bar{\alpha}^M m_{it} + \bar{\alpha}^L l_{it} + \bar{\alpha}^K k_{it} + \bar{\beta} s_{it} m_{it}).$$

Our generalized method of moments (GMM) procedure assumes that plant-level innovations to productivity, $\zeta_{it}(\Omega)$, do not correlate with the predetermined choices of inputs. To recover $\zeta_{it}(\Omega)$, we assume that productivity for any set of parameters, $\bar{\omega}_{it}(\Omega)$, follows a first order Markov process. Thus, we can approximate the productivity process with the following function:

$$\bar{\omega}_{it}(\Omega) = \gamma_0 + \gamma_1 \bar{\omega}_{i,t-1}(\Omega).$$

³⁵Gandhi et al (2020) argue that identifying the flexible input's output elasticity from the condition is difficult because it is not entirely clear if flexible input demand is high because of the high productivity or the high output elasticity.

From the equation above, we can recover the innovation to productivity, $\zeta_{it}(\Omega)$, for a given set of parameters. Since the productivity term, $\bar{\omega}_{it}(\Omega)$, can be correlated with the current choices of variable inputs, l_{it} and m_{it} , but it is not correlated with the fixed input, k_{it} , the innovation to productivity, $\zeta_{it}(\Omega)$, will not be correlated with $\mathbf{Y}_{it} = \{k_{it}, l_{i,t-1}, m_{i,t-1}, \text{ and } s_{i,t-1}m_{i,t-1}\}$. Thus, we use the following moment condition:

$$E[\zeta_{it}(\Omega)\mathbf{Y}_{it}] = 0 \quad (18)$$

and search for the optimal combination of the parameters by minimizing the sum of the moments (and driving it as close as possible to zero) using the standard weighting procedure for plausible values of Ω .

Table A2 reports output elasticity estimates. The output elasticities of intermediate inputs, labor, and capital are on average 0.59, 0.36, and 0.08, respectively. These estimates match closely with the cost shares reported in Table 2.

III. The Melitz-Polanec Decomposition

The findings from the United States (i.e., De Loecker et al, 2020; Autor et al, 2020) suggest that the increasing market shares of superstar firms are responsible for the rise of aggregate markups. During Japan's lost decades after the burst of the bubble economy in 1991, however, such competitive selection did not occur in Japanese manufacturing. Caballero et al (2008) argue that the widespread practice of Japanese banks' continued lending on nonperforming loans in the 1990s kept unproductive firms alive and distorted competition.

To better understand the trend of markups over the period we consider, we follow Melitz and Polanec (2015) and decompose the 1987-2007 change in the weighted markup from the entire sample ($\Delta\mu^V$) into the between and within effects, and also include the exit and entry effects:

$$\Delta\mu^V = \Delta cov_s + \Delta\bar{\mu}_s^V + w_{x,87}(\mu_{s,87}^V - \mu_{x,87}^V) + w_{e,07}(\mu_{e,07}^V - \mu_{s,07}^V) \quad (19)$$

where Δcov_s is the change in the covariance between markup (μ_{it}^V) and the revenue share (w_{it}) in the survivor sample, $\Delta\bar{\mu}_s^V$ is the change in the mean of plant-level markups across survivors, w_{xt} (w_{et}) is the aggregate revenue share of the exiters (the entrants) in the full sample in year t , and

μ_{xt}^V , μ_{et}^V , and μ_{st}^V are the weighted mean markups of the exiters, the entrants, and the survivors, respectively.

The between effect (Δcov_s) is the change in the covariance, and a higher positive value indicates that survivors with higher markups gain higher market shares. The within effect ($\Delta \bar{\mu}_s^V$) is the change in the mean markups for survivors. The exit effect is the difference between the weighted mean markups of survivors versus exiters in 1987; and, the entry effect is the difference between the weighted mean markups of entrants versus survivors in 2007.

Table A4 reports the results from markups when intermediate inputs and labor are variable inputs. Since the mean and covariance are sensitive to outlier values, we drop the top and bottom additional one percent of the markup distributions for each year. Thus, the aggregate change reported in columns (4) and (6) in Table 3 does not perfectly match the change in Table A4. In the first row of the table, the left-hand side variable, the change in the weighted markup from the full sample of plants over the period of 1987 to 2007 declined by 0.7 percentage points. The aggregate markup was depressed by the between effect (i.e., a 3.5 percentage point decline), caused by the markets allocated to plants with lower values of markups. The within effect and the entry of new plants contributed to a 0.8 percentage point and a 1.7 percentage point increase, respectively.

The second to fifth rows in Table A4 report the results when we examine the weighted markups at the industry level. The markup in the automobile industry increased by 2.4 percentage points, and the exit of low markup plants contributed to the increase. The markup in the textile industry increased by 8.4 percentage points, and the between effect contributed the most. The markup in the communication equipment industry decreased by 9.2 percentage points, and the exit of high markup plants was responsible for the decline.

References

- [1] Akerberg, D., K. Caves, and G. Frazer, “Identification Properties of Recent Production Function Estimators,” *Econometrica*, 83, 2411-2451, 2015.
- [2] Asker, J., A. Collard-Wexler, and J. De Loecker, “Dynamic Inputs and Resource (Mis)Allocation,” *Journal of Political Economy*, 122(5), 1013-1063, 2014.
- [3] Atkeson, A, and A. Burstein, “Pricing-to-Market, Trade Costs, and International Relative Prices,” *American Economic Review*, 98(5), 1998-2031, 2008.
- [4] Autor, D., D. Dorn, L.F. Katz, C. Patterson, and J. Van Reenen, “The Fall of the Labor Share and the Rise of Superstar Firms,” *Quarterly Journal of Economics*, 135(2), 645-709, 2020.
- [5] Barkai, S., “Declining Labor and Capital Shares,” *Journal of Finance*, 75(5), 2421-2463, 2020.
- [6] Basu, S., “Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence,” *Journal of Economic Perspectives*, 33(3), 3-22, 2019.
- [7] Bills, M., “The Cyclical Behavior of Marginal Cost and Price,” *American Economic Review*, 77(5), 838-855, 1987.
- [8] Bond, S., A. Hashemi, G. Kaplan, and P. Zoch, “Some Unpleasant Markup Arithmetic: Production Function Elasticities and their Estimation from Production Data,” *Journal of Monetary Economics*, 121, 1-14, 2021.
- [9] Caballero, R.J., T. Hoshi, and A.K. Kashyap, “Zombie Lending and Depressed Restructuring in Japan,” *American Economic Review*, 98(5), 1943-77, 2008.
- [10] Caballero, R.J., and R.K. Lyons, “External Effects in U.S. Procyclical Productivity,” *Journal of Monetary Economics*, 29, 209-225, 1992.
- [11] Collard-Wexler, A., “Demand Fluctuations in the Ready-mix Concrete Industry,” *Econometrica*, 81(3), 1003-1037, 2013.
- [12] Cooper, R.W., and J.C. Haltiwanger, “On the Nature of Capital Adjustment Costs,” *Review of Economic Studies*, 73(3), 611-633, 2006.

- [13] De Loecker, J., “Comment on (Un)pleasant ... by Bond et al (2020),” mimeo, 2020.
- [14] De Loecker, J. and J. Eeckhout, “Global Market Power,” NBER Working Paper #24768, 2020.
- [15] De Loecker, J., J. Eeckhout, and G. Unger, “The Rise of Market Power and the Macroeconomic Implications,” *Quarterly Journal of Economics*, 135(2), 561-644, 2020.
- [16] De Loecker, J., P.K. Goldberg, A.K. Khandelwal, and N. Pavcnik, “Prices, Markups, and Trade Reform,” *Econometrica*, 84(2), 445-510, 2016.
- [17] De Loecker, J., and F. Warzynski, “Markups and Firm-Level Export Status,” *American Economic Review*, 102(6), 2437-2471, 2012.
- [18] De Ridder, M., B. Grassi, and G. Morzenti, “The Hitchhiker’s Guide to Markup Estimation,” mimeo, 2021.
- [19] Diewert, W.E., and K.J. Fox, “On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups,” *Journal of Econometrics*, 145, 174-193, 2008.
- [20] Doraszelski, U., and J. Jaumandreu, “Measuring the Bias of Technological Change,” *Journal of Political Economy*, 126(3), 1027-1084, 2018.
- [21] Feenstra, R.C., and D.E. Weinstein, “Globalization, Markups and US Welfare,” *Journal of Political Economy*, 125(4), 1040-74, 2017.
- [22] Foster, L.S., J.C. Haltiwanger, and C.J. Krizan, “Aggregate Productivity Growth: Lessons from Microeconomic Evidence,” in “New Developments in Productivity Analysis,” University of Chicago Press, 303-372, 2001.
- [23] Foster, L.S., J.C. Haltiwanger, and C. Tuttle, “Rising Markups or Changing Technology?” NBER Working Paper 30491, 2022.
- [24] Foster, L., J.C. Haltiwanger, and C. Syverson, “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, 98(1), 394-425, 2008.
- [25] Fukao, K., S. Hamagata, T. Inui, K. Ito, H.N. Kwon, T. Makino, T. Miyagawa, Y. Nakanishi, and J. Tokui, “Estimation Procedures and TFP Analysis of the JIP Database 2006,” RIETI Discussion Paper Series 07-E-003, 2007.

- [26] Fukao, K., and K. Ito, “Output Quality, Skill Intensity, and Factor Contents of Trade: An Empirical Analysis Based on Microdata of the Census of Manufactures,” RIETI Discussion Paper Series 10-E-028, 2010.
- [27] Gandhi, A., S. Navarro, and D.A. Rivers, “On the Identification of Gross Output Production Functions,” *Journal of Political Economy*, 128(8), 2973-3016, 2020.
- [28] Giannetti, M., and A. Simonov, “On the Real Effects of Bank Bailouts: Micro Evidence from Japan,” *American Economic Journal: Macroeconomics*, 5(1), 135-167, 2013.
- [29] Hall, R.E., “The Relation between Price and Marginal Cost in U.S. Industry,” *Journal of Political Economy*, 96(5), 921-47, 1988.
- [30] Hall, R.E., and D.W. Jorgenson, “Tax Policy and Investment Behavior,” *American Economic Review*, 57(3), 391-414, 1967.
- [31] Hashimoto, M., and J. Raisian, “Employment Tenure and Earnings Profiles in Japan and the United States,” *American Economic Review*, 75(4), 721-735, 1985.
- [32] Hopenhayn, H.A., “Entry, Exit, and firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60(5), 1127-1150, 1992.
- [33] Hsieh, C.T., and P.J. Klenow, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(4), 1403-1448, 2009.
- [34] Jorgenson, D.W, and Z. Griliches, “The Explanation of Productivity Change,” *Review of Economic Studies*, 34(3), 249-283, 1967.
- [35] Kambayashi, R., and T. Kato, “Long-Term Employment and Job Security over the Past 25 Years: A Comparative Study of Japan and the United States,” *Industrial Labor Relations Review*, 70(2): 359-394, 2017.
- [36] Karabarbounis, L., and B. Neiman, “Accounting for Factorless Income,” NBER Working Paper #24404, 2018.
- [37] Kiyota, K., T. Nakajima, and K.G. Nishimura, “Measurement of the Market Power of Firms: the Japanese Case in the 1990s,” *Industrial and Corporate Change*, 18(3), 381-414, 2009.

- [38] Kugler M., and E. Verhoogen, “Prices, Plant Size, and Product Quality,” *Review of Economic Studies*, 79(1), 307-339, 2012.
- [39] Levinsohn, J. and A. Petrin, “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, 70(2), 317-342, 2003.
- [40] Melitz, M.J., and S. Polanec, “Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit,” *RAND Journal of Economics*, 46(2), 362-375, 2015.
- [41] Olley, G.S., and A. Pakes, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64(6), 1263-1297, 1996.
- [42] Raval, D., “Testing the Production Approach to Markup Estimation,” *Review of Economic Studies*, 2023.
- [43] Roeger, W., “Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for U.S. Manufacturing,” *Journal of Political Economy*, 103(2), 316-30, 1995.
- [44] Syverson, C., “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy*, 112(6), 1181-1222, 2004.
- [45] Syverson, C., “Macroeconomics and Market Power,” *Journal of Economic Perspectives*, 33(3), 23-43, 2019.
- [46] Traina, J., “Is Aggregate Market Power Increasing? Production Trends Using Financial Statements,” University of Chicago Booth School of Business Stigler Center for the Study of the Economy and the State Working Paper #17, 2018.

Figures and Tables

Table 1. Summary statistics

	Mean			s.d.		
	1987	2007	Δ	1987	2007	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
log real revenue	11.625	12.068	0.444	1.317	1.338	0.021
log real spending on intermediates	10.936	11.247	0.311	1.577	1.572	-0.005
log labor	4.374	4.429	0.055	0.803	0.815	0.012
log real capital stock	10.200	10.869	0.669	1.598	1.679	0.081

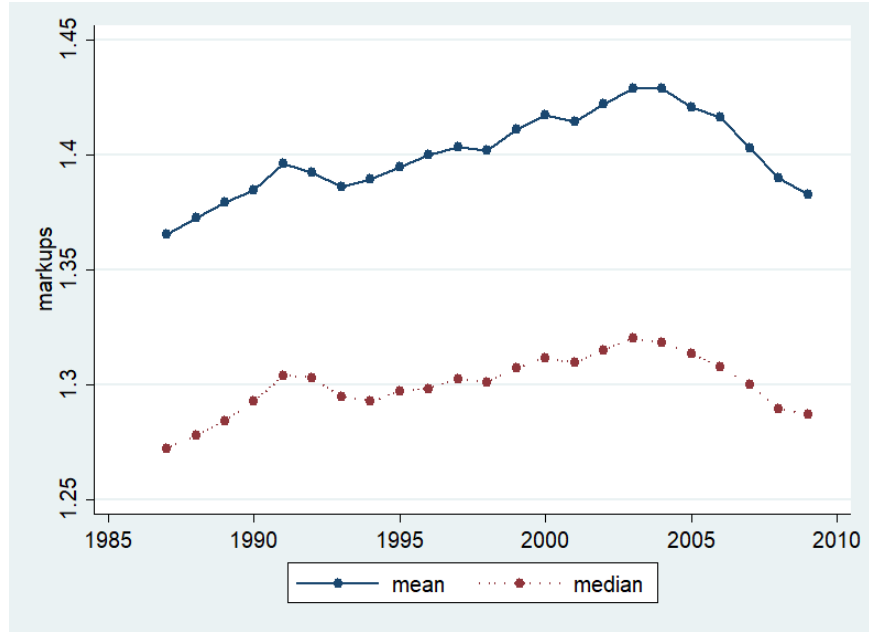
Notes: (1) We have 44,817 observations in 1987 and 39,641 observations in 2007. (2) See Appendix I for the development strategy of real capital stock. (3) The unit is billion Japanese Yen in 2000 for real revenues, real spending on intermediates, and real capital stock.

Table 2. Cost shares of six types of inputs

	Mean		s.d.	
	1987	2007	1987	2007
	(1)	(2)	(3)	(4)
Intermediate inputs (M)				
Materials	0.499	0.515	0.245	0.235
Fuels	0.012	0.015	0.022	0.031
Electricity	0.022	0.023	0.030	0.027
Outsourcing	0.110	0.102	0.140	0.141
Factor inputs				
Labor (L)	0.303	0.288	0.185	0.173
Capital (K)	0.053	0.057	0.050	0.056

Notes: We use Jorgenson and Griliches (1967) to compute capital cost. See Appendix I.

Figure 1. Mean and median markups in Japan's manufacturing



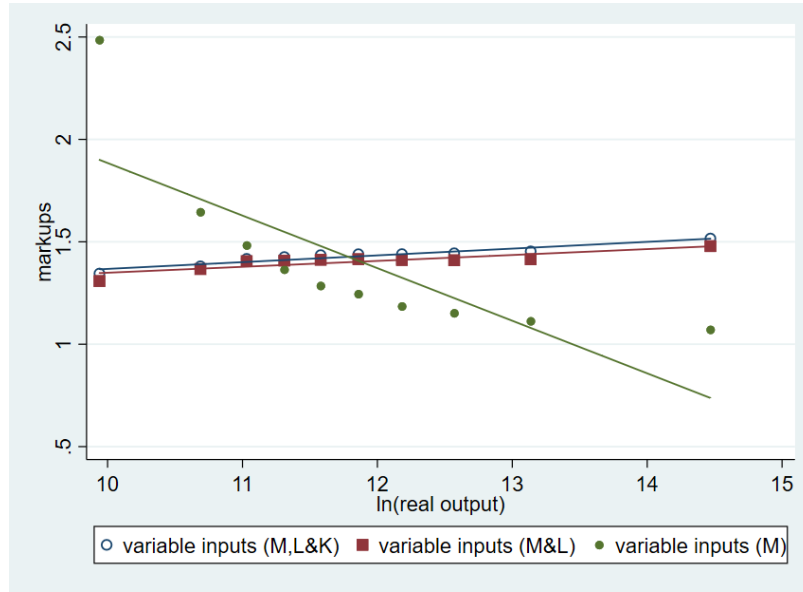
Notes: (1) We use intermediate inputs and labor as variable inputs to compute markups. See equation (5) for the definition of markups. (2) We drop the top and bottom 1% of yearly observations as outliers.

Table 3. Variable inputs and markup estimates

Variable inputs:	Intermediates (M)		M and labor (L)		M, L and capital (K)	
	1987	2007	1987	2007	1987	2007
	(1)	(2)	(3)	(4)	(5)	(6)
Weighted mean	1.095	1.028	1.436	1.435	1.473	1.474
Distribution						
Mean	1.354	1.357	1.369	1.408	1.395	1.426
s.d.	0.921	0.931	0.378	0.440	0.370	0.421
Percentile						
10th	0.708	0.705	1.021	1.007	1.055	1.040
25th	0.862	0.863	1.128	1.137	1.162	1.170
50th	1.095	1.105	1.275	1.302	1.305	1.328
75th	1.515	1.528	1.497	1.552	1.524	1.567
90th	2.970	2.870	2.122	2.259	2.110	2.244

Notes: (1) See equation (5) for the definition of markups. (2) We use plant-level revenues as a weight to compute the weighted means. (3) We drop the top and bottom 1% of yearly observations for each markup measure as outliers.

Figure 2. Plant size and three types of markups



Notes: (1) We use binned scatterplots and fitted lines to visualize associations between log (real output) and markups in 1997. Here, real output is computed from the revenue divided by the industry-level output deflator. (2) See Tables 4 and 5 for conditional correlations between these two variables.

Table 4. The determinants of the yearly changes in markups

Dependent variable:	$\Delta \ln(\text{markup})$					
Price and quantity data:	Product-level price			Industry price deflator		
Variable inputs:	M	M and L	M, L and K	M	M and L	M, L and K
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(\text{output price})$	0.928*** (0.010)	0.912*** (0.010)	0.927*** (0.010)	0.942*** (0.017)	0.762*** (0.039)	0.794*** (0.036)
$\Delta \ln(\text{real output})$	-0.450*** (0.025)	-0.070*** (0.018)	-0.014 (0.018)	-0.459*** (0.022)	-0.001 (0.014)	0.049*** (0.013)
$\Delta \ln(\text{TFP})$	1.378*** (0.026)	0.989*** (0.021)	0.941*** (0.019)	1.404*** (0.026)	0.861*** (0.027)	0.820*** (0.024)
$\Delta \ln(\text{variable input price})$	-0.830*** (0.022)	-0.891*** (0.033)	-0.900*** (0.030)	-0.813*** (0.017)	-0.698*** (0.043)	-0.717*** (0.043)
Observations	57,325	57,325	57,325	870,935	870,935	870,935
R-squared	0.827	0.891	0.898	0.825	0.825	0.835

Notes: (1) We use the single-product plant data for the first panel (columns (1)-(3)) and the entire sample for the second panel (columns (4)-(6)). (2) Output price, real output, and TFP are based on product-level prices and quantities for the first panel, and they are computed from industry-level deflators for the second panel. (3) Standard errors that are clustered at the industry level are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence level, respectively.

Table 5. The determinants of log markups

Dependent variable:	ln(markup)					
Price and quantity data:	Product-level price			Industry price deflator		
Variable inputs:	M	M and L	M, L and K	M	M and L	M, L and K
	(1)	(2)	(3)	(4)	(5)	(6)
ln(output price)	0.717*** (0.068)	0.784*** (0.035)	0.796*** (0.035)	1.024*** (0.046)	0.584*** (0.034)	0.606*** (0.036)
ln(real output)	-0.171*** (0.018)	0.001 (0.006)	-0.004 (0.005)	-0.202*** (0.021)	0.007 (0.005)	0.007 (0.004)
ln(TFP)	0.876*** (0.047)	0.774*** (0.033)	0.793*** (0.031)	1.086*** (0.017)	0.659*** (0.028)	0.661*** (0.026)
ln(variable input price)	-0.769*** (0.050)	-0.714*** (0.072)	-0.727*** (0.075)	-1.019*** (0.112)	-0.301*** (0.056)	-0.313*** (0.056)
Industry-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67,125	67,125	67,125	987,299	987,299	987,299
R-squared	0.567	0.778	0.807	0.694	0.674	0.714

Notes: See Table 4.

Figure 3. Input adjustments over the dot-com bubble collapse



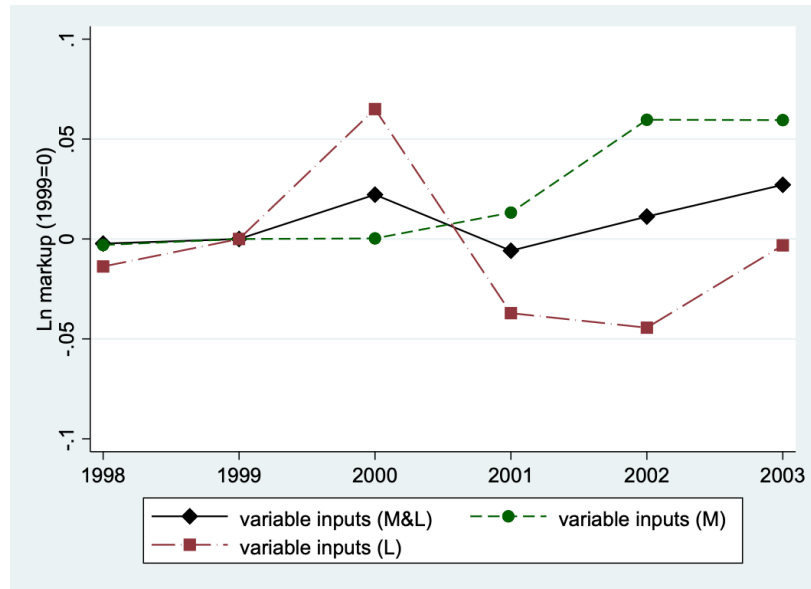
Notes: (1) We use 591 establishments in the semiconductor industry. (2) The variables are normalized to zero in 1999 at the establishment level. We report the means within the industry.

Table 6. Input adjustments over the dot-com bubble collapse

	Mean		
	2000 (1)	2002 (2)	Δ (3)
Intermediate inputs (M)	11.501	11.172	-0.329
Materials	10.793	10.441	-0.352
Fuels	6.194	6.071	-0.116
Electricity	8.411	8.309	-0.098
Outsourcing	9.799	9.368	-0.490
Factor inputs			
Labor (L)	5.081	4.933	-0.149
Capital (K)	10.978	10.808	-0.169

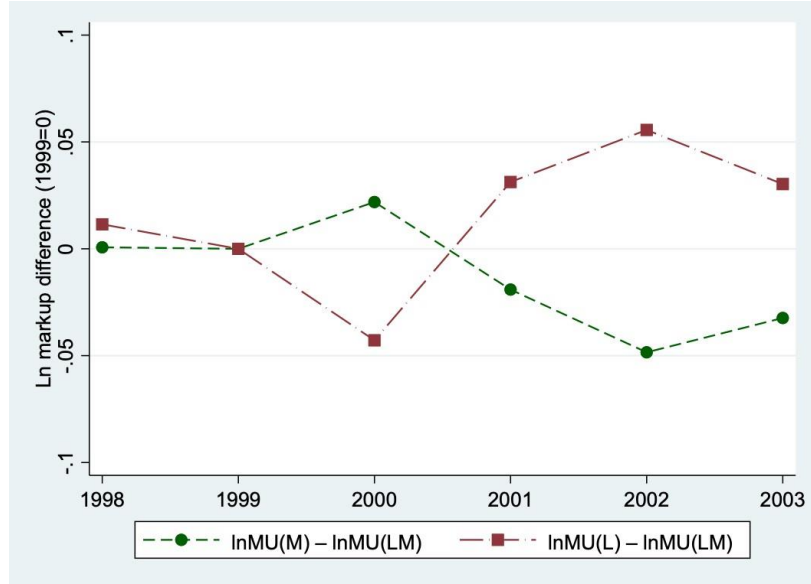
Notes: (1) We use 591 establishments in the semiconductor industry. (2) All variables are expressed in logarithms. (3) See Figure 2 for the long-run adjustments for selected inputs.

Figure 4. Changes in markups in the semiconductor industry over the dot-com bubble



Notes: (1) We use 591 establishments in the semiconductor industry. (2) The variables are logs of markups normalized to zero in 1999 at the establishment level. We report the means within the industry.

Figure 5. Changes in marginal costs in the semiconductor industry over the dot-com bubble



Notes: (1) We use 591 establishments in the semiconductor industry. (2) The variable is log of markup using material (labor) as a variable input subtracted by log of markup using material and labor as variable inputs. The variables are normalized to zero in 1999 at the establishment level. We report the means within the industry.

Table 7. Markups and market shares

Dependent variable:	ln(market share)					
Market definition:	Industry-level markets			Industry × prefecture markets		
Variable inputs:	M	M and L	M, L and K	M	M and L	M, L and K
	(1)	(2)	(3)	(4)	(5)	(6)
ln(markup)	-0.975*** (0.060)	0.579*** (0.087)	0.648*** (0.081)	-0.758*** (0.044)	0.602*** (0.055)	0.621*** (0.048)
Control variables						
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	987,299	987,299	987,299	987,299	987,299	987,299
R-squared	0.407	0.324	0.326	0.256	0.223	0.223

Notes: (1) We define the sum of shipments in each industry as a market in the first panel (columns (1)-(3)) and the sum of shipments in each industry in a prefecture as a market in the second panel (columns (4)-(6)). (2) Standard errors that are clustered at the industry level are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence level, respectively.

Appendix

Table A1. Coverage of the sample

	Mfg Census (1)	JIP database (2)	Coverage (3) = (1)/(2)
Revenue (billion ¥)			
1987	180,984	265,175	0.683
2007	245,878	338,719	0.726
Labor (1,000)			
1987	6,003	13,848	0.433
2007	5,476	11,061	0.495
Capital (billion ¥)			
1987	62,457	138,744	0.450
2007	109,228	234,726	0.465

Notes: (1) See Fukao et al (2007) for the development of the JIP database. (2) See Appendix I for the development strategy of capital stock.

Table A2. Summary statistics of estimated output elasticities

	Mean (1)	s.d. (2)	Min (3)	Max (4)
Intermediate inputs (α_M)	0.591	0.059	0.466	0.897
Labor (α_L)	0.363	0.074	0.102	0.488
Capital (α_K)	0.082	0.026	0.019	0.207

Notes: See Appendix II for our estimation strategy.

Table A3. Summary statistics of markups, prices, costs, and productivity

	Mean			s.d.		
	1987	2007	Δ	1987	2007	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
Log output price						
ln(output deflator)	0.048	-0.010	-0.059	0.214	0.205	-0.010
ln(product price)	2.603	3.027	0.425	2.214	2.183	-0.031
Log real output						
ln(revenue/output deflator)	11.625	12.068	0.444	1.317	1.338	0.021
ln(product physical quantity)	8.711	8.739	0.029	2.739	2.711	-0.027
Log TFP						
ln(revenue TFP)	2.732	2.877	0.145	0.556	0.605	0.049
ln(quantity TFP)	0.442	0.027	-0.415	2.279	2.219	-0.060
Log variable input price						
Intermediates	-0.008	0.078	0.086	0.101	0.138	0.037
Intermediates and labor	-0.196	-0.040	0.155	0.261	0.223	-0.038
Intermediates, labor and capital	-0.178	-0.047	0.131	0.248	0.210	-0.038
Log market share						
Industry $\times$ prefecture markets	-4.666	-4.505	0.161	1.712	1.731	0.019
Industry-level markets	-8.148	-8.056	0.092	1.474	1.466	-0.007

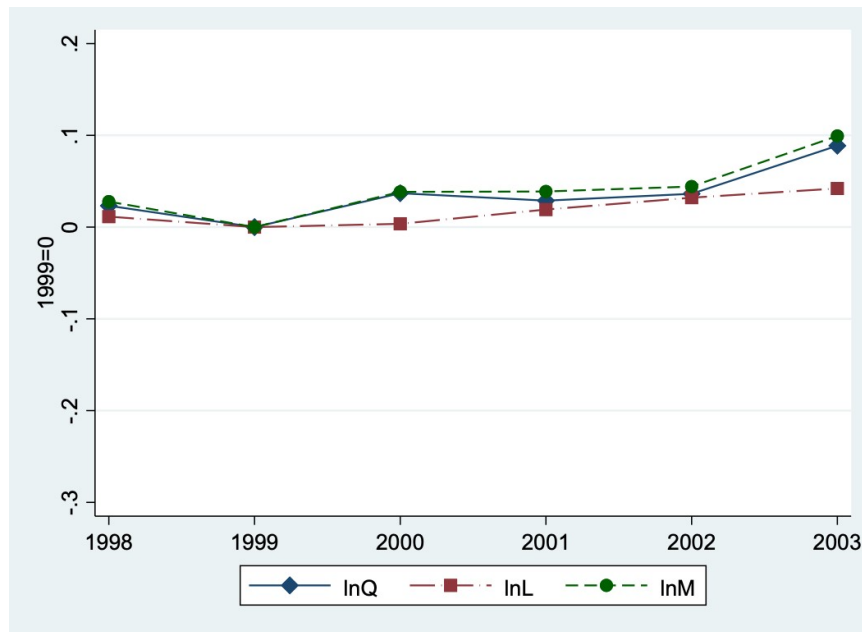
Notes: We have 44,817 observations in 1987 and 39,641 observations in 2007 for output deflator, revenue/output deflator, revenue TFP, unit variable input costs, and market shares. We use 3,348 observations in 1987 and 2,110 observations in 2007 for product price, physical quantity, and quantity TFP from single-product plants.

Table A4. The Melitz and Polanec decomposition (1987-2007)

	Survivors			Exit and entry			All plants
	Between	Within	(1)+(2)	Exit	Entry	(4)+(5)	(3)+(6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All industries	-0.035	0.008	-0.027	0.003	0.017	0.020	-0.007
Industries							
Automobile	-0.003	0.003	0.000	0.061	-0.038	0.023	0.024
Apparel and textile	0.060	0.014	0.074	0.021	-0.011	0.010	0.084
Communication	-0.107	0.096	-0.011	-0.139	0.057	-0.082	-0.092
Precision machinery	0.002	0.048	0.050	-0.005	0.013	0.008	0.058

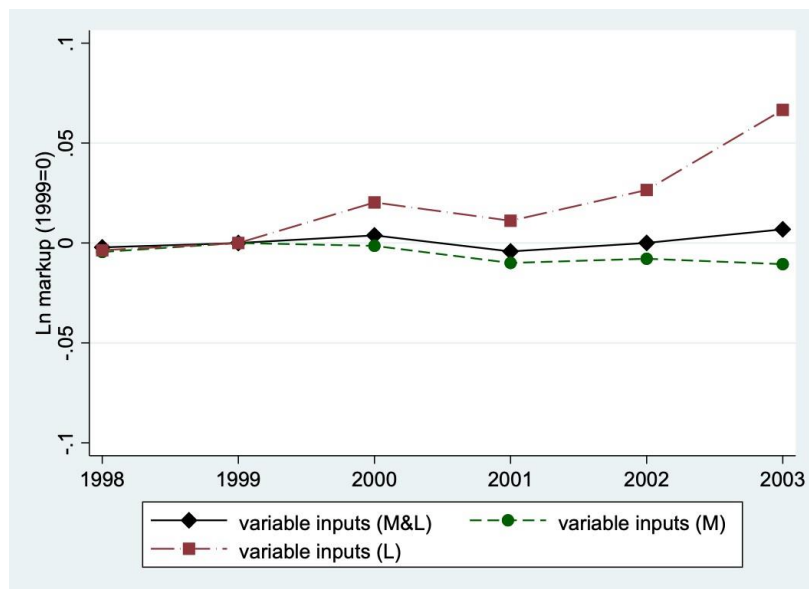
Note: (1) See equation (19). (2) We use markups using intermediate inputs and labor and use plant-level revenues as weights. (3) We additionally drop the top and bottom 1% of observations for each year.

Figure A1. Input adjustments over the dot-com bubble collapse



Notes: (1) We use 1,619 establishments in the automotive industry. (2) The variables are normalized to zero in 1999 at the establishment level. We report the means within the industry.

Figure A2. Changes in markups over the dot-com bubble



Notes: (1) We use 1,619 establishments in the automotive industry. (2) The variables are logs of markups normalized to zero in 1999 at the establishment level. We report the means within the industry.