Industrial Policy, Location Choice, and Firm Performance in High-Tech Manufacturing*

Ran Zhuo† Audrey Tiew‡

August 14, 2026

Abstract

We study how industrial policies promoting domestic self-reliance affect semiconductor manufacturers' incentives to build and continue operating production capacity across locations in the contract manufacturing market for semiconductor logic chips. Combining worldwide facility-level investment data with global contract manufacturing orders, we estimate a model of contracting for these chips from 2004 to 2015 that recovers firm-, location-, and buyer-proximity-specific cost advantages. In a case study of TSMC's location choice for a large, advanced facility, we find that initial fixed-cost subsidies comparable, as a share of investment, to those provided under the CHIPS and Science Act more than offset the additional $1.2 billion cost of building in the U.S. However, relocating to the U.S. results in $1.6–$1.8 billion lower profits over our study period than locating a comparable facility in Taiwan, because higher estimated location-specific marginal costs outweigh the benefits of proximity to U.S. buyers. These losses are concentrated in later years, as entry into maturing technology generations erodes margins and demand shifts toward Asian buyers. Front-loaded fixed-cost subsidies therefore satisfy the firm's initial participation constraint but not its continuation constraints as the technology matures. Among the policy instruments we study, only tariffs preserve the firm's incentive to remain in the U.S., but they impose welfare losses on U.S. buyers that exceed the firm's gains.

*We are very grateful to Shane Greenstein, Myrto Kalouptsidi, Robin Lee, Ariel Pakes, and Elie Tamer for their time, patience, and insightful advice. We also thank William Cao, Ashley Chang, Taiyi Cheng, Renee James, Willy Shih, H.-S. Philip Wong, and David Yoffie for thoughtful discussions about the semiconductor industry and industrial policy. We thank David M. Byrne, Ed Hall, Jessica Mueller, and Shungo Saito for assistance with data-related inquiries. We are grateful to Judy Chevalier, Ying Fan, Jessie Handbury, Panle Jia, Brian Wu, and participants at the Harvard Industrial Organization Workshop; the Ross Technology & Operations Brown Bag and Hosmer-Adderley Faculty Lunch; the Columbia Business School IEOR-DRO Seminar; the INFORMS Annual Meeting 2025; the WashU Boeing Center Research Seminar; WBEC 2026; and the NBER Digitization Winter 2026 meeting for helpful suggestions. We thank the Doctoral Office at Harvard Business School, the Technology & Operations Department and the Dean's Office at the Ross School of Business for financial assistance. We thank Steven Kim for excellent research assistance. All errors are our own. Contributions follow the CRediT taxonomy (National Information Standards Organization, 2022), with roles designated as lead, equal, or supporting. **Ran Zhuo:** Conceptualization (lead), Methodology (supporting), Formal Analysis (equal), Data Curation (lead), Software (lead), Writing - original draft (lead), Writing - review & editing (equal), Investigation (lead), Validation (equal), Visualization (lead), Funding Acquisition (lead). **Audrey Tiew:** Methodology (lead) , Formal Analysis (equal), Validation (equal), Writing - review & editing (equal), Conceptualization (supporting), Investigation (supporting), Software (supporting), Visualization (supporting), Writing - original draft (supporting).

†Contact: ranzhuo@umich.edu. Affiliation: Ross School of Business, University of Michigan.

‡Contact: audreytiew@nyu.edu. Affiliation: Department of Economics, New York University.

1 Introduction

High-tech manufacturing in industries such as biotechnology, electric vehicles, and semiconductors is characterized by rapid technological turnover, substantial fixed-cost investments, and significant economies of scale. Consider semiconductor manufacturing: building and equipping a leading-edge facility cost less than $1 billion before the 2000s, around $5 billion in the 2010s, and over $10 billion in the 2020s (Taiwan Semiconductor Manufacturing Company, 2020). The fundamental economics of these industries increasingly clash with recent national policies in the U.S. and China that emphasize domestic self-reliance in key high-tech sectors. These policies aim to boost domestic production through generous investment subsidies, tax incentives, export controls, and tariffs (see, e.g., The State Council of the People's Republic of China, 2015; U.S. Senate Committee on Commerce, Science, and Transportation, 2022a; York and Durante, 2025), thereby changing manufacturers' incentives for capacity investment across regions.

We ask how these policies affect firms' incentives to initially invest in, and to later maintain, semiconductor manufacturing capacity in different locations. Policies operate through two channels. The first is investment costs: the up-front construction and equipment outlays required to build or upgrade capacity, which become sunk once incurred. The second is flow profits: the stream of profits a facility earns each period, net of its recurring fixed and variable costs, which evolves over the facility's life as its technology matures. Promoting persistent domestic production requires addressing both channels. We measure the first channel with an investment cost function learned from worldwide facility-level investment data, and the second with a structural model of global contract manufacturing that recovers cost advantages specific to firms, to locations, and to the proximity between manufacturers and their buyers.

Our central finding is that these two channels have opposite time profiles. The additional investment cost of locating production in the U.S. is front-loaded and relatively modest. For the large advanced facility constructed between 2000 and 2004 that is the focus of our policy counterfactual, we estimate this additional cost at $1.2 billion—an amount more than covered by a lump-sum subsidy comparable, as a share of investment, to those provided under the CHIPS and Science Act. The flow-profit penalty, by contrast, is larger and back-loaded. Locating the same facility in the U.S. reduces its owner's profits by $1.6–$1.8 billion over our 2004–2015 simulation period, when the facility is in production, with the losses concentrated late in the

facility's life.

Although our empirical analysis focuses on a recent historical period, the economic forces driving these results—including the predictable evolution of market structure as technologies mature and the associated shifts in demand—remain central to the industry today. The results therefore speak directly to current industrial policy. They reflect the diffusion of production capabilities across manufacturers over a technology's life cycle. At any point in time, only a small number of manufacturers can produce at the technology frontier, whereas progressively more firms become capable of producing older generations. Margins on a given technology therefore erode predictably as the frontier advances and entry into that generation accumulates. Changes in the geography of demand reinforce this pattern. Demand for advanced chips comes disproportionately from U.S. buyers, allowing a U.S. facility to benefit from buyer-proximity cost advantages early in its life. As the technology matures, however, demand shifts increasingly toward Asian buyers, turning this initial proximity advantage into a disadvantage. The opportunity cost of locating in the U.S. is therefore smallest when the facility operates at the frontier and largest after its technology has matured and the initial investment has become sunk. Industrial policy, by contrast, is largely front-loaded. This creates a mismatch between the time profile of the location penalty and that of the policy instrument. Up-front fixed-investment subsidies may therefore satisfy a firm's initial participation constraint but do not address its incentives to continue operating in the U.S. as the technology matures.

Our empirical setting is the contract manufacturing market in the semiconductor industry, in which manufacturers produce semiconductor logic chips on behalf of other companies, primarily fabless designers that lack their own fabrication facilities. Their clients range from leading chip designers like Apple and Nvidia to buyers of legacy chips for defense and medical applications. The market is highly concentrated, with Taiwan Semiconductor Manufacturing Company (TSMC) alone accounting for more than half of the revenue earned by the top eight manufacturers, and it is a central target of recent industrial policies in both countries, from CHIPS and Science Act subsidies and tax credits to export controls and tariffs (Section 2.4 provides details). The number of concurrent policies makes it difficult to isolate the effects of individual interventions using reduced-form methods like event studies, motivating a structural approach that can explore counterfactual policy scenarios.

We combine a comprehensive, worldwide sample of facility-level capacity investments observed quarterly from 1995 to 2015 with a representative sample of global contract manufacturing orders observed quarterly from 2004 to 2015, supplemented by publicly available data on firm revenues, gross margins, technology and region revenue shares, and capacity utilization. We use these data to estimate the cost parameters that determine market equilibrium during the 2004 to 2015 period, and our counterfactual analysis focuses on the same window. This timeframe precedes the major policy shifts that began in late 2014, so the industry was more likely to be in the stable equilibrium that our estimation requires.

Our model endogenizes manufacturer profits and how they vary with the number of competitors and with capacity, using a stylized bargaining framework in which buyers solicit bids, quarter by quarter, from all manufacturers with available capacity for a given technology generation. Capacity investment across locations and the evolution of demand over time are held fixed. Negotiation is captured by a second-price procurement auction: the winning manufacturer prices just above the second-best competitor's effective marginal cost, so margins reflect relative costs and the number of capable bidders. The model is specified to capture the tradeoff between firm-specific and location-specific cost advantages: each manufacturer's marginal cost includes common technology costs, a firm-specific component, regional cost differences, capacity scale effects, and idiosyncratic shocks, and buyers gain extra value from contracting with a manufacturer in their own region. In our main specification, buyers have perfectly inelastic demand for chip quantities, reflecting downstream rigidity.[1]

We estimate these parameters using the method of simulated moments, matching moments constructed from different data sources. Because our orders data report only the region in which the fabricating facility is located, not the identities of buyers or sellers, order-level price variation identifies the common generation- and period-specific costs, while variation in firm revenues and gross margins from quarterly reports identifies firm- and location-specific costs, buyer-location parameters, and fixed costs. Crucially, combining the capacity data with firm revenues is what separates firm cost advantages from location cost advantages: firms operate fabs in multiple regions, so shifts in a firm's revenue when its regional capacity mix changes identify the location component separately from the firm component.

Our parameter estimates confirm key economic intuitions about the global contract semi-

[1]We test this assumption empirically and suggest a model extension to relax it.

conductor manufacturing market. The estimates indicate that more advanced technologies are more costly to produce, while production costs decline over time as technologies mature. Firm- and region-specific cost differences further reveal substantial advantages for major players such as TSMC. On average, TSMC's firm-specific efficiency lowers its marginal cost by an amount equivalent to 13.96% of the common generation- and period-specific cost relative to the average manufacturer in Taiwan. In addition, locating a fab in Taiwan rather than North America lowers marginal cost by an amount equivalent to 9.43% of the common generation- and period-specific cost. The model also identifies a significant cost benefit from geographic proximity: when buyers contract with manufacturers in the same region, marginal cost is lower by an amount equivalent to 7.96% of the common generation- and period-specific cost, on average. On the demand side, the estimates show that North American buyers dominate demand for cutting-edge technologies, but their share declines over time and as technologies mature, while mainland Chinese buyers become increasingly prominent.

We then combine the model estimates with our investment cost function in a detailed case study of TSMC's payoffs to locating a large manufacturing facility in the U.S. versus its home region under various industrial policies. The ideal experiment, TSMC's Arizona investment announced in 2020, falls outside our data and was developed amid frequent policy changes. We instead study TSMC Fab 14, a facility within our sample period whose investment profile closely resembles Arizona's development timeline, phased strategy, and adoption of cutting-edge technology with preplanned upgrades (Section 2.5 details the comparison). In our counterfactuals, we vary only this facility's location—Taiwan or the U.S.—while holding fixed the capacity and technology upgrade trajectories of this facility and of all other facilities in the market. The location comparison can thus be read as a partial derivative of the firm's investment problem with respect to the location choice, and the policy scenarios as cross-partial effects that measure how each policy alters that comparison.

We find that export controls and Chinese tax exemptions substantially reduce TSMC's profits but leave the relative attractiveness of the two locations nearly unchanged: the profit penalty from locating in the U.S. remains between $1.6 billion and $1.8 billion across these policy scenarios. In the case of export controls, this is because the buyers excluded by the restrictions would have contracted with TSMC regardless of where the fab was located. Tariffs are the

exception because they operate directly on the same margin that drives the location penalty. By repricing contracts according to the locations of buyers and manufacturers, tariffs improve the competitiveness of a U.S. facility precisely where it is weakest: in mature technologies serving price-sensitive demand. Under our calibration, introducing tariffs reverses the location comparison, from a \$1.6 billion loss to a \$5.9 billion gain, making the U.S. location privately optimal for the firm.

This realignment, however, comes at a substantial cost to U.S. buyers. When the facility remains in Taiwan, tariffs generate \$48.9 billion in welfare losses for U.S. buyers, of which \$44.0 billion reflects tariff payments and \$4.9 billion reflects increased contracting with less cost-efficient U.S. manufacturers. When the facility is relocated to the U.S., U.S. buyers source more frequently from domestic production, reducing tariff payments to \$36.0 billion but increasing the welfare loss from contracting with the less cost-efficient U.S. fabs to \$11.8 billion. As a result, relocation generates an additional \$7.0 billion in welfare losses from production cost inefficiency, exceeding the firm's \$5.9 billion gain from moving production to the U.S. Policies that strengthen private incentives for domestic semiconductor manufacturing may therefore impose costs on domestic chip users that exceed the corresponding gains to manufacturers.

Our paper's primary contribution is to the economics of industrial policy (see Juhász et al., 2024, for a review). A large recent empirical literature uses historical and quasi-experimental variation to estimate the effects of policies that protect or promote targeted industries (e.g., Juhász, 2018; Criscuolo et al., 2019; Lane, 2025; Choi and Levchenko, 2025). A smaller quantitative literature uses model-based approaches to evaluate industrial policies and compare alternative policy instruments (e.g., Baldwin and Krugman, 1988; Kalouptsidi, 2018; Barwick et al., 2025).[2] Our paper belongs to this latter tradition. A model-based approach is particularly useful in our setting because several policies operate simultaneously, while some of the policies we study are prospective. The model allows us to evaluate and compare individual policy instruments and combinations of instruments, including those for which no realized policy variation is available.

Our results also connect to recent work on the incidence and production effects of U.S. trade policy. Consistent with evidence that recent U.S. tariffs have been passed through sub-

[2]Baldwin and Krugman (1988) is an early quantitative analysis of strategic trade policy, while the latter papers estimate structural models using microdata.

stantially to domestic purchasers (Fajgelbaum et al., 2020) and can induce firms to relocate production (Flaaen et al., 2020), we find that tariffs can shift semiconductor manufacturing geographically while imposing substantial costs on domestic chip buyers. We study these forces in a concentrated business-to-business market in which production requires large, sunk, and location-specific capacity investments. This setting highlights a distinct policy issue that has received less attention: the time profile of an industrial-policy instrument may be misaligned with that of the location disadvantage it is intended to offset.

We also contribute to the growing literature on industrial policy in semiconductors (see Bown and Wang, 2024, for an overview). A large policy-oriented literature from government agencies, industry associations, think tanks, and market-intelligence and consulting firms has documented the geographic concentration of semiconductor production and evaluated policies intended to alter it (e.g., Friedberg and Boustany Jr, 2020; Semiconductor Industry Association, 2021; Bateman, 2022; U.S. Government Accountability Office, 2022; Kim and VerWey, 2019; Varas and Varadarajan, 2020; U.S. Chamber of Commerce, 2021; Varas et al., 2020). These studies range from qualitative institutional assessments to quantitatively detailed exercises combining industry data, engineering cost benchmarks, and scenario-based forecasts. For example, Varas et al. (2020) projects that a $50 billion U.S. manufacturing subsidy could lead to the construction of 19 new fabrication facilities and capture an additional 24% of the global manufacturing capacity needed between 2020 and 2030. Our analysis complements these exercises by estimating the behavioral primitives underlying firms' location and investment responses. In particular, rather than positing how changes in production costs translate into capacity relocation, we use observed capacity and transactions to recover manufacturer- and location-specific production costs and buyer-proximity frictions, and then compare the effects of alternative policies on firm profits across production locations.

A growing academic literature examines particular semiconductor policies, including export controls, technological decoupling, and the CHIPS Act (Jacobs et al., 2022; Han et al., 2024; Erten et al., 2025). The model-based studies closest to ours are Goldberg et al. (2024) and Miao (2024). Goldberg et al. (2024) documents industrial policy across semiconductor-producing countries and estimates substantial firm-product learning-by-doing, economies of scope, and cross-border learning spillovers. Miao (2024) focuses on the technology race among the frontier

manufacturers and estimates a dynamic oligopoly model of technology adoption under trade disruption risk, while abstracting from capacity investment and technology investment in more mature technologies. Our contribution relative to these studies is to focus on a central policy margin in current semiconductor industrial policy: the geographic relocation of manufacturing capacity. We study manufacturers' incentives to relocate production across countries, compare profits under alternative production locations, identify which policy instruments can induce relocation, and quantify the resulting implications for domestic chip buyers.

Methodologically, our paper contributes to the empirical industrial organization literature on investment in concentrated, capital-intensive industries (e.g., Goettler and Gordon, 2011; Ryan, 2012; Kalouptsidi, 2014; Igami, 2017; Igami and Uetake, 2020; Thurk, 2022). Our key contribution is to connect the short-run allocation of buyer orders to firms' long-run, location-specific capacity decisions. Combining buyer–manufacturer transactions with facility-level capacity and financial data allows us to separately identify the production-cost penalty associated with locating a given technology in different regions and the buyer-proximity frictions that shape contracting outcomes. We then carry these primitives into firms' capacity-investment decisions. This separation is particularly important in semiconductors because capacity is long-lived while the economics of a process technology change substantially as it matures. Unlike a standard quality-ladder setting in which older technologies are displaced by the frontier, mature semiconductor technologies continue to serve distinct downstream applications and buyers. What changes is the set of firms capable of producing them and the margins that the technologies command. Our framework therefore identifies how the return to a location-specific capacity investment evolves over the technology life cycle and makes it possible to evaluate industrial policies whose duration may be short relative to the life of the capital they induce.

The remainder of the paper proceeds as follows. Section 2 describes the contract semiconductor manufacturing market, the maturation of technology generations, the policy environment, and the case study that anchors our analysis. Section 3 describes the data. Section 4 presents the model of contracting, Section 5 the estimation method, and Section 6 the estimation results. Section 7 presents the policy counterfactuals, Section 8 discusses the limitations of our approach, and Section 9 concludes.

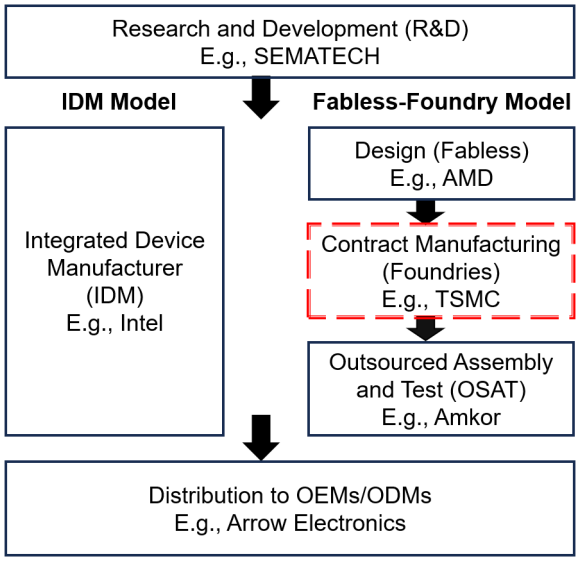
2 Empirical Setting

This section describes the contract manufacturing market, the product and its technology generations, how entry reshapes competition as technologies mature, and the policy environment. It concludes by introducing the case study that anchors our analysis: TSMC's choice of where to locate a large fabrication facility.

2.1 Contract Manufacturing in Semiconductor Industry Value Chain

The semiconductor industry has a complex value chain characterized by two major production models and a number of distinct types of firms. The first is the Integrated Device Manufacturer (IDM) model. IDMs design, manufacture, assemble, and test semiconductor chips in-house and then sell the finished chips to downstream Original Equipment Manufacturers (OEMs) and Original Design Manufacturers (ODMs). The second is the fabless-foundry model, in which fabless firms design chips and outsource their manufacture to contract manufacturers. When a chip designer orders chips from a contract manufacturer, the contract specifies the quantity, technical specifications, price, and delivery schedule. "Pure-play" foundries emerged with the founding of TSMC in 1987 and specialize exclusively in manufacturing chips ordered by their clients. Figure 1 provides a simplified illustration of where contract manufacturing is situated within the semiconductor industry value chain.

Fig. 1: Contract manufacturing in semiconductor industry value chain



Note: Adapted from Figure 3 of Semiconductor Industry Association (2016).

The distinction between the two models is not strict. IDMs may outsource production when

their internal capacity is constrained, and they may manufacture chips for external customers when they have excess capacity or dedicated foundry operations, as in the case of Samsung. One important difference between IDM and foundry manufacturing lies in the product mix. IDM fabs are typically optimized for a relatively limited range of products sold by the IDM itself, whereas foundries manufacture a much more diverse set of products for multiple customers. For example, an Intel fab typically produces Intel microprocessors, while a TSMC fab may manufacture microprocessors, systems-on-chip used in smartphones, and AI chips.

Our analysis of contract manufacturing revenue of top contract manufacturers shows that two patterns stand out. First, pure-play foundries dominate the contract manufacturing market. Throughout 2005–2017 where we have data on contract manufacturing revenue for both top pure-play foundries and IDMs, IDMs have only accounted for 2–11% of the revenue of the top eight players in this market. Second, the market is oligopolistic and is dominated by a few big players. One firm, TSMC, accounts for 52–61% of the revenue among the top eight contract manufacturers over the period 2005–2017. TSMC has been by far the largest and the most important player in this market. Other major players include United Microelectronics Corporation (UMC) in Taiwan, Semiconductor Manufacturing International Corporation (SMIC) in China, GlobalFoundries in the U.S., and Samsung in South Korea.[3]

2.2 The Product and Manufacturing Technology

The basic unit of production in semiconductor manufacturing is a silicon wafer, on the surface of which layers of integrated circuits are fabricated. Chip designers provide manufacturers with blueprints of their chip designs, and manufacturers use light together with a series of complex physicochemical processes to transfer these intricate circuit patterns onto the wafer, layer by layer. Each wafer contains many individual chips, which are subsequently cut from the wafer and packaged. For a given chip design, different manufacturers can produce chips based on the same specifications. The resulting products are therefore largely homogeneous from the chip designer's perspective, and manufacturers compete primarily on production efficiency and cost. In particular, higher yields of functioning chips, lower defect rates, and lower manufacturing costs

[3]UMC and SMIC consistently rank among the top eight during 2005–2017, accounting for 9–21% and 5–8% of top-eight revenue, respectively. The remaining positions fluctuate due to entry, mergers, and other industry dynamics. GlobalFoundries, formed in 2009 from the divestiture of AMD's manufacturing operations, accounts for 7–14% of top-eight revenue between 2009 and 2017. Samsung officially launched its dedicated foundry business in 2005 and accounts for 5–11% of top-eight revenue between 2010 and 2017.

translate into greater price competitiveness and, ultimately, larger market shares. The most prevalent manufacturing technology is the complementary metal-oxide-semiconductor (CMOS) process, which is the focus of our study and accounts for 92% of the orders in our representative demand data.[4]

The evolution of manufacturing technology follows a numerical nanometer scale. For example, different generations are referred to as 90nm, 65nm, 40nm, and so on. Smaller technology nodes represent more advanced manufacturing capabilities than larger ones: a smaller node produces smaller integrated circuits fitted more tightly into the same surface area on the wafer, and delivers higher performance, more energy efficiency, and more cost efficiency for the chips. Through the 28nm generation, which entered commercial production around 2012, node names measure a specific physical dimension of the chip. For newer generations, node names are naming conventions that manufacturers use to signal substantive improvements over the previous generation rather than exact physical measurements. Appendix D.8 discusses the evolution of node naming conventions and how we map node names to technology generations.

2.3 Technology Maturation, Entry, and Competition

Semiconductor manufacturing is a prime example of a high-tech industry characterized by substantial fixed costs and rapid technological turnover. Roughly every two years, consistent with Moore's Law, a new generation of manufacturing technology is introduced and the technology frontier advances. As a result, each existing generation moves progressively farther from the frontier. As shown in Figure 2A, at any point in time only a small number of manufacturers operate at the technology frontier, while progressively more firms are capable of producing generations farther behind the frontier. The number of manufacturers using a given generation also follows a predictable life cycle: entry is concentrated in the first four to six years after introduction, after which the number of producers gradually declines as firms retire increasingly mature technologies. The competitive environment surrounding a manufacturing technology therefore evolves predictably as the technology ages.

Prices also evolve predictably over this maturation process. At a given point in time, more advanced generations generally command higher prices, while the price of a given generation tends to decline as it matures, as shown in Figure 2B. This decline can reflect both learning

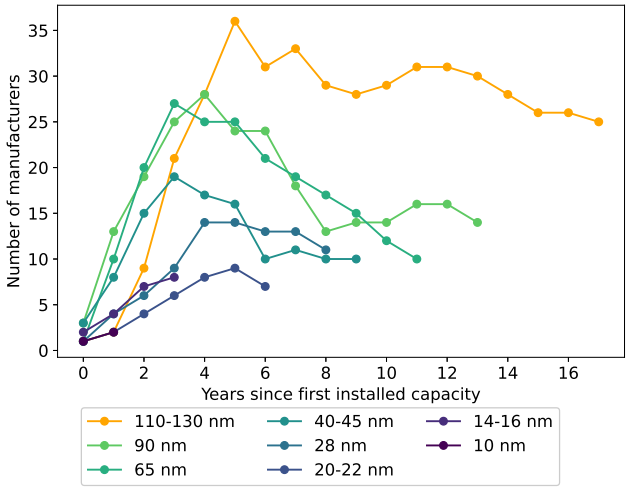
[4]The next most prevalent technology, the BiCMOS process, accounts for less than 4% of the orders in our demand data.

by doing, as manufacturers become more efficient at producing a given generation and reduce production costs, and increasing competition, as the diffusion of production capabilities erodes the market concentration that initially supports high margins.

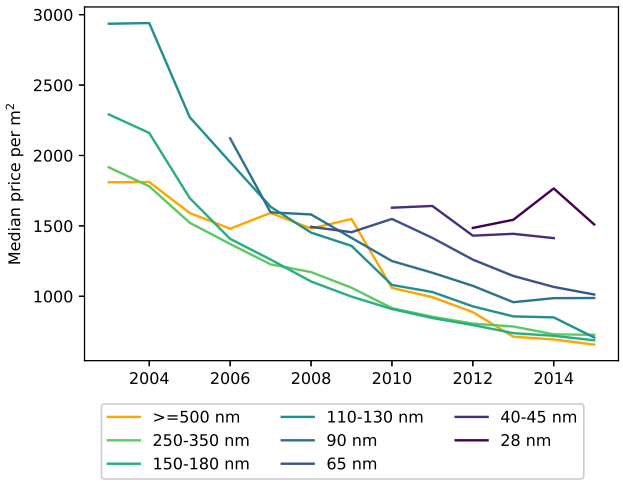
Revenue shifts across generations in a similarly systematic way. As new manufacturing technologies are introduced, they gain revenue share at the expense of older generations, as shown in Figure 2C. This pattern reflects both demand and supply forces. On the demand side, high-value chips such as microprocessors require advanced manufacturing technologies, causing orders to migrate toward newer generations as they become available. On the supply side, when a leading manufacturer such as TSMC first introduces a new generation, few competitors are initially capable of offering the same technology, supporting high prices and margins. Over time, additional manufacturers enter the generation, intensifying price competition and reducing its prices, revenues, and margins.

Fig. 2: Technology turnover and scale economies of semiconductor manufacturing

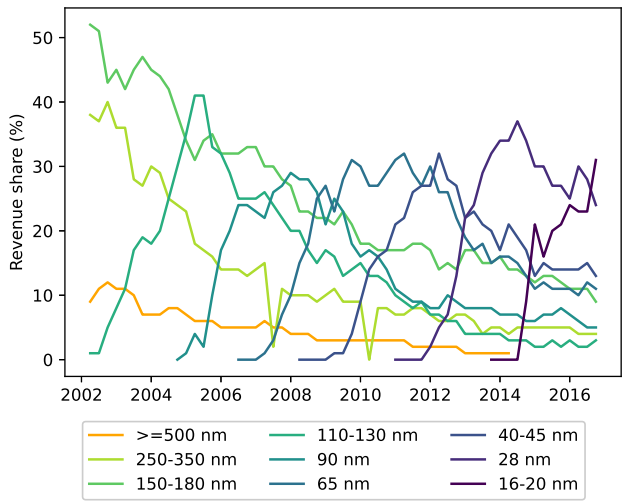
(A) Number of manufacturers by technology



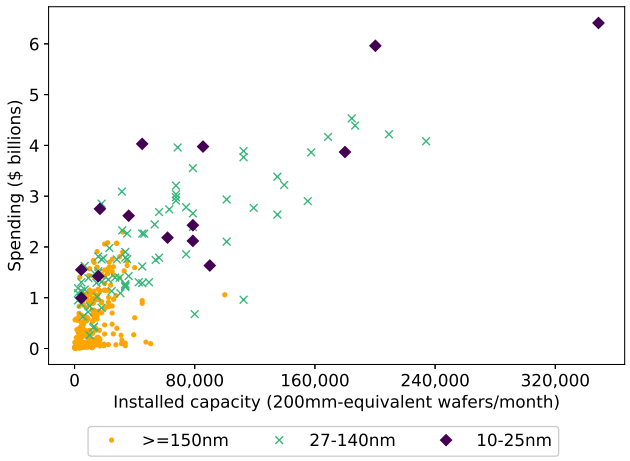
(B) Median price by technology



(C) TSMC revenue shares by technology



(D) New fab construction and equipment spending



Notes: Data source for Figure 2A is SEMI. See Table 3 for summary statistics. The number of manufacturers includes both pure-play foundries and IDMs. Curves for more advanced technology generations are progressively shorter because these generations enter later in the sample and are therefore truncated by the end of the sample period. Data source for Figure 2B is GSA. See Table 2 for the definition of price per m^2. Data source for Figure 2C is TSMC quarterly reports. Figure 2D covers new fab construction worldwide between 1995 and 2015. To calculate the installed capacity, wafers of different sizes have been converted to their equivalents in wafers of 200mm in diameter, a standard practice in the semiconductor industry for capacity calculation. Spending figures are inflation-adjusted, with a base value of 100 assigned to January 2015.

At the same time, advancing manufacturing technologies have become increasingly expensive. Manufacturers therefore tend to build fewer but larger fabs for newer technologies in order to exploit scale economies, as shown in Figure 2D. The industry thus combines a technology frontier that advances on a roughly two-year cadence with progressively larger investments required to remain near that frontier.

Together, these patterns motivate the hypothesis that the profits a fab can earn from a

given technology evolve systematically as that technology recedes from the frontier and entry accumulates. A fab that begins production at the frontier therefore earns its highest margins early in the technology life cycle, when capable competitors are few, and progressively lower margins as the technology matures and additional competitors enter—precisely after the initial construction and equipment expenditures have become sunk. Whether it is optimal for a manufacturer to continue operating, upgrade, or disinvest from a fab in a given location therefore depends on how far the fab's technology has receded from the frontier and on how policies alter its payoffs at different stages of this maturation process.

2.4 Policy Background

Recent policies in the U.S. and China on the semiconductor industry differ from past government support, which historically focused on R&D and tax concessions (OECD, 2019), in both the much larger scale of direct subsidies and their explicit focus on domestic manufacturing. The four instruments considered in our counterfactuals are the main tools of this new policy wave. In the U.S., the CHIPS and Science Act of 2022 provides $39 billion in direct subsidies to build, expand, and modernize domestic fabs, as well as a 25% investment tax credit (U.S. Senate Committee on Commerce, Science, and Transportation, 2022a,b). The U.S. has also imposed export controls restricting sales of advanced chips and manufacturing equipment to Chinese entities (Khan, 2020; Industry and Security Bureau, 2022). On China's side, the first and second phases of the "Big Fund," launched in 2014 and 2019, respectively, provided $50 billion in state equity financing directed primarily toward manufacturing capacity (Semiconductor Industry Association, 2020, 2021), complemented by corporate tax exemptions for manufacturers operating advanced technology nodes (Ministry of Finance of the People's Republic of China et al., 2020). Tariffs on semiconductors have also been under active discussion in the U.S. since 2025 (York and Durante, 2025). Appendix A.1 reviews these policies, as well as smaller programs in other countries, in greater detail.

While the U.S. semiconductor industry largely supported and welcomed the subsidies (Semiconductor Industry Association, 2022), U.S. geographic restrictions on investment in China and sales of chips to China have been the subject of substantial industry lobbying, as the industry collectively earns roughly one-third of its revenue from China. Chipmakers have expressed concerns that overly restrictive measures could disrupt supply chains and hamper their global

competitiveness (Swanson, 2023; Hayashi and Fitch, 2023). The Semiconductor Industry Association (SIA), which represents more than 95% of the U.S. semiconductor industry by revenue and nearly two-thirds of non-U.S. chip firms, states that "the semiconductor industry is truly global, and access to global markets is critical for U.S. firms to sustain high levels of investment in R&D and capital expenditure" (Semiconductor Industry Association, 2021). Our conversations with industry experts similarly indicate that economies of scale and access to business from the global market are crucial to the health and growth of the industry. These potentially conflicting incentives between national governments and industry motivate our study.

It is important to note that the policy landscape for semiconductor manufacturing has been highly uncertain and continuously evolving since the early 2020s, and is likely to remain so over the next three to five years. Recent developments illustrate this volatility. On China's side, the third phase of the Big Fund was announced in 2024 (Reuters, 2024), with an additional $48 billion in financing. In the U.S., the 2025 Budget Reconciliation Act increased the investment tax credit (ITC) for semiconductor investment from 25% to 35% for property placed in service after 2025 and for which physical construction begins before the end of 2026 (CNBC, 2025). In addition, discussions of semiconductor tariffs, which have persisted since early 2025, are likely to continue, with proposed rates reaching 25% or higher (York and Durante, 2025) and actual rates ranging between 10–14% over 2025–2026 (McClelland and Wong, 2026). This continued policy flux not only creates a highly uncertain environment for firms, but also makes it difficult to isolate the effects of individual policies using an event-study framework. It therefore motivates our model-based structural estimation and counterfactual approach, which allows us to evaluate policy bundles and firms' strategic responses under alternative policy scenarios, including policies implemented outside our sample period.

2.5 TSMC Fab Location Case Study

To anchor our examination of how different policy bundles influence major manufacturers' location investment choices and flow profits, we study a detailed case of a major semiconductor firm's decision between locating a large fabrication facility in its home region or in the U.S. amid shifting industrial policies.

Ideally, we would directly analyze TSMC's decision to build fabs in Arizona, initially announced in 2020 with subsequent additional investments announced in the following years, as

the Arizona fab cluster represents the key strategic response by the world's leading contract manufacturer to rising geopolitical tensions and U.S. policy initiatives aimed at attracting advanced semiconductor production. However, this case falls outside our sample period of 2004–2015, and we lack comprehensive data on firms' investment behavior, capacity, and demand conditions during this more recent period. Additionally, the Arizona facilities were developed and brought into production amid frequent policy changes, making it difficult to reliably infer key cost and competitive parameters, as the market was likely operating outside of a stable equilibrium.

To address these challenges, we identify comparable investment projects within our sample period that feature observable market conditions and outcomes under a relatively stable policy environment. This allows us to more credibly recover the key structural parameters of interest. We find the development of TSMC Fab 14 particularly relevant, as it closely resembles the investment profile of TSMC Arizona. Table 1 compares the Arizona and Fab 14 facilities, revealing TSMC's consistent fab investment approach over time while also highlighting important similarities and differences. Both facilities followed a similar four-year timeline from construction announcement to production start and launched with TSMC's cutting-edge technologies of their respective eras (4nm for Arizona in 2024; 130nm for Fab 14 in 2004). Each project adopted a phased construction and expansion strategy and demonstrated planned/actual technology node migrations, though TSMC Arizona begins at a significantly more advanced node. Fab 14 eventually transitioned to a mature node facility, typical for TSMC fabs after multiple upgrades, while TSMC Arizona's long-term trajectory remains unclear. The most notable distinction lies in the Arizona project's substantial government subsidies, precisely the main policy variation our counterfactuals examine, and higher total investment, reflecting the growing complexity and cost of advanced-node manufacturing.

To analyze the likely impact of different industrial policy bundles, including capacity-building subsidies, tax incentives, export controls, and tariffs, we estimate key competitive parameters for the years 2004–2015 and simulate counterfactual outcomes for TSMC, comparing scenarios in which Fab 14, originally a cutting-edge facility in Taiwan, either remains in Taiwan or is hypothetically relocated to the U.S. These counterfactuals provide insight into how such policies might influence the outcomes of investments like TSMC Arizona.

Table 1: Comparative Development of TSMC Facilities

TSMC Arizona	TSMC Fab 14 in Taiwan
Timeline: • 2020: First fab announced, construction began • 2022: Second fab annouced • 2024: First fab production began in Q4 (4nm), third fab announced • 2028 (planned): Second fab production (3nm) • 2030 (planned): Third fab production (2nm) **Technology Position:** • Launched with 4nm (one generation behind most advanced technology at time of production) • Planned migration to 3nm and 2nm nodes **Development Approach:** • Three-phase expansion strategy • 4-year timeline from announcement to production **Investment & Subsidies:** • Total investment: \$65B across three phases • Government subsidy: \$6.6B from U.S. Department of Commerce (2024)	**Timeline:** • 2000: Phase 1 construction began • 2002: Phase 2 construction began • 2004: Phase 1 production began in Q2 (130nm) • 2007: Phase 2 production began in Q2 (55nm) **Technology Position:** • Launched with 130nm (most advanced at time of production) • Multiple technology migrations: 130nm $\rightarrow$ 90nm $\rightarrow$ 80nm $\rightarrow$ 65nm $\rightarrow$ 55nm $\rightarrow$ 45nm $\rightarrow$ 40nm • Eventually settled as mature node facility (40nm since 2008) **Development Approach:** • Two-phase expansion strategy • 4-year timeline from announcement to production **Investment:** • Total investment: \$6.6B b/w 2000–2015 • No significant government subsidies noted

3 Data

Our unique data combine a range of rich industry data sources, including both a representative sample of worldwide contract manufacturing orders and a comprehensive sample of worldwide fab-level capacity investment. For more details about sample and variable construction as well as additional ancillary data, please see Appendix D.

3.1 Demand

Information on wafer orders comes from a proprietary database of wafer purchases by chip designers from contract manufacturers, collected by the Global Semiconductor Alliance (GSA), an industry organization originally founded to support fabless semiconductor firms worldwide but that has since expanded its membership to the entire semiconductor ecosystem. The dataset consists of individual responses to the GSA's quarterly Wafer Fabrication & Back-End Pricing Survey, detailing the orders that chip designers purchased during 2003Q4 to 2015Q3.

Prior research deem the data a representative sample of wafer orders produced by the contract manufacturing sector worldwide (Byrne et al., 2017; Thurk, 2022; Goldberg et al., 2024). For each order, the data report the quantity of wafers ordered, the price paid per wafer, the location of the fab, the manufacturing technology, wafer size, and the number of layers fabricated. This information allows us to examine order characteristics by manufacturing location. Although this level of product detail is remarkable, the GSA data have important limitations (Byrne et al., 2017). We only observe the country or region in which each fab was located but have no information on the producing firm or the buyer due to anonymization. Table 2 shows summary statistics of this dataset. Because wafers differ in diameter and in the number of layers fabricated, the table reports quantities and prices in three unit conventions: raw wafers, 200mm-equivalent wafers, which is the standard industry convention, and layer-adjusted square meters, which is the unit in which our model and estimation measure quantities and prices.

3.2 Fab Capacity and Investment

Information on fab capacity and investment comes from a proprietary database collected by SEMI, the global industry association serving the semiconductor manufacturing supply chain. The dataset we obtained covers the universe of fabs worldwide on the quarterly basis for 1995–2015. For each fab-quarter, the data report the installed capacity, manufacturing technology, wafer size, any ongoing investment (either in construction or equipping the fab), the type of investment (e.g., constructing a new fab or upgrading an existing fab), and investment amount. The data also provides a rich set of fab characteristics, including the location, ownership, and whether the fab manufactured specific types of products (e.g. micropressors) or was a dedicated foundry for contract manufacturing of a wide range of products. This dataset allow us to examine the capacity and investment decisions of contract manufacturers.

Table 2: Summary statistics of wafer orders, 2003Q4–2015Q3

	Count	Mean	Std	Min	Max
Quantity (# wafers)	11927	2704.59	9977.55	1	53000
Quantity (# 200mm-equivalent wafers)	11927	3530.20	14946.03	1.00	530000.00
Quantity (m^2)	11927	3462.05	16741.43	0.44	466212.34
Price/wafer	11927	1578.58	1302.25	100.00	10340.00
Price/200mm-equivalent wafer	11927	1192.93	651.34	152.00	8000.00
Price/m^2	11927	1437.13	1626.96	212.21	90541.48
# of layers	11927	27.66	8.47	1	55
Wafer size (mm)	11927	217.10	46.06	100	300
Technology generation	11927	10.83	2.01	6	17
Fab location					
China	11927	0.05	0.23	0	1
North America	11927	0.06	0.23	0	1
ROW	11927	0.20	0.40	0	1
Taiwan	11927	0.69	0.46	0	1

Note: Quantity in 200mm-equivalent wafers= #wafers×(surface area of wafer ordered/ surface area of 200mm wafer). Measuring wafer quantities in 200mm equivalents is standard practice in this industry. Quantity in m^2 = #wafers×#of layers×surface area/wafer. Price/200mm-equivalent wafer=price/wafer÷(surface area of wafer ordered/ surface area of 200mm wafer). Price/m^2 =price/wafer÷(#of layers×surface area/wafer). Technology nodes (in nm) are converted into integer values representing technology generations. The conversion process is detailed in Appendix D.8. All monetary values are in USD and are not inflation-adjusted.

However, it is important to note that while our data on fab investment is detailed, it lacks a breakdown of the investment amount into manufacturers' own out-of-pocket spending (through reinvestment of revenue or financing) and subsidies and other incentives received from local and national governments. These benefits are not always transparent (OECD, 2019), and our data provider has confirmed the difficulty in collecting comprehensive information about government incentives in fab investments. Table 3 shows the summary statistics of this dataset, with capacity and capacity changes reported in the same three unit conventions as the wafer orders.

3.3 Industry and Firm Performance

We supplemented our two proprietary data sources by collecting additional information on industry and firm performance from a variety of publicly available data sources. We gathered annual contract manufacturing revenue figures for top contract manufacturers from data tables released by IC Insights, a semiconductor industry intelligence firm. This dataset spans top pure-play foundries for 2000–2004 and includes both top pure-play foundries and IDMs for 2005–2017.

Table 3: Summary statistics of fab capacity and investment, 1995Q1–2015Q4

	Count	Mean	Std	Min	Max
Fab capacity					
Capacity (wafers/month)	35293	23405.79	23540.31	25	245000
Capacity (200mm-equiv wafers/month)	35293	21698.17	38232.27	20	528750
Capacity (m^2/month)	35293	20627.40	51879.50	2.52	773459.70
IDM fab	35293	0.79	0.41	0	1
Wafer size (mm)	35293	170.69	54.25	75	300
Technology generation	35293	8.65	3.28	3	19
Fab location					
China	35293	0.10	0.30	0	1
North America	35293	0.24	0.43	0	1
ROW	35293	0.56	0.50	0	1
Taiwan	35293	0.10	0.29	0	1
Fab investment projects					
Investment amount (millions)	1185	395.28	665.24	1.50	5725
Construction investment amount (millions)	1185	62.47	165.19	0	1540
Equipment investment amount (millions)	1185	332.81	531.00	0	4210
Project duration (quarters)	1185	9.09	7.01	1	58
Construction duration (quarters)	1185	1.44	2.27	0	13
Change in tech generation	1185	2.93	4.60	-3	18
Change in capacity (wafers/month)	1185	12856.94	19814.20	-15000	175000
Change in capacity (200mm-equiv wafers/month)	1185	17658.98	36264.54	-6750	393750
Change in capacity (m^2/month)	1185	20320.27	45874.66	-1487.28	575980.63
Investment type					
New construction	1185	0.26	0.44	0	1
Tech upgrade	1185	0.10	0.30	0	1
Capacity expansion	1185	0.33	0.47	0	1
Upgrade and expansion	1185	0.18	0.38	0	1
Maintenance	1185	0.13	0.34	0	1

Note: Capacity in m^2/month= wafers/month×predicted avg layers/wafer× surface area/wafer. Avg layers/wafer predicted by linear regression of observed avg layers/wafer for each tech generation on the tech generation ($R^2 = 0.984$). Observations with extremely old technologies >3300nm dropped due to negative predicted layers. Construction duration corresponds to the observed quarters of construction spending. Project duration equals the greater of the observed quarters of construction and equipment spending or the quarters needed for the fab to reach maximum capacity within a year after the investment ends. Negative capacity change may occur when tech is upgraded. Negative tech change may occur when capacity is expanded. For new constructions, initial capacity and technology generation are set to zero, so their changes equal the constructed facility's capacity and technology generation. 86 investment projects dropped due to missing capacity information. 24 investment projects dropped where both tech change and capacity change in 200mm equivalent were nonpositive. This happened due to various reasons, such as significant changes in product mix, fabs having multiple lines that invested in older technology than their most advanced line, etc. While we hope to investigate this further, the scarcity of these observations and the granularity of data required prevent us from doing so. Projects starting before 1995Q1 or ending after 2015Q4 are excluded due to incomplete investment data outside this range, as the raw data only records investment amounts in the quarters they occurred. Appendix D details the sample construction and excluded facilities. The rest of notes of Table 2 apply.

Additionally, we collected quarterly firm performance metrics, including contract manufacturing revenue, gross margin, R&D spending, and capacity utilization, from quarterly reports of publicly listed contract manufacturers. One limitation of this dataset is that IDMs do not separately report their contract manufacturing revenue, so our data exclusively covers pure-play foundries. Our dataset includes information for TSMC, UMC, SMIC, and four other pure-play

foundries. For leading contract manufacturers like TSMC, UMC, and SMIC, their quarterly reports provide further insights into the revenue shares of different generations of manufacturing technologies and across various geographical regions. We also gather this type of information whenever it is available.

To assess the overall macroeconomic demand for semiconductors, we utilized publicly available data from the World Semiconductor Trade Statistics (WSTS) on aggregated semiconductor sales. These figures are based on net billings between semiconductor firms and their end customers, authorized distributors, and divisions or subsidiaries that manufacture end products. WSTS, a non-profit mutual benefit corporation, serves the global semiconductor industry by collecting monthly revenue data from its members and distributing it in aggregate form back to them. This data series is extensively used by the semiconductor industry to gauge overall industry demand.

4 Model

We endogenize manufacturer profits and how they change with the number of competitors and capacity constraints using a stylized model of bargaining, where each buyer engages in multilateral bargaining with multiple manufacturers. This model reflects key pricing drivers in the contract semiconductor manufacturing industry, including production costs and competitive pressure from rivals (Acquired, 2025). Specifically, we assume buyers solicit bids from all manufacturers with available capacity for the required technology generation, with capacity fixed within each period. This structure effectively transforms the negotiation into a second-price procurement auction, a modeling of negotiation similar to that in Allen et al. (2019) and Miller et al. (2025). Each manufacturer's marginal cost decreases as installed capacity increases. Our model parsimoniously captures how firms achieve higher markups when fewer competitors produce the same generation of technology and when they have larger capacities.

A Preferences and Cost Function

Let i denote an individual chip designer that demands contract manufacturing capacity for the generation of manufacturing technology g in period t. Buyer i's maximum willingness to pay per unit quantity of generation g in period t is denoted as v_{igt}. When more than 1 manufacturer has available capacity and participates in bidding, the buyer maximizes the following utility

function by choosing which contract manufacturer, denoted as j, to use:

$$\max_{j \in \mathcal{J}_{igt}} v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\}) - p^j_{igt}. \tag{1}$$

J_{igt} denotes the set of contract manufacturers that have available capacity and participate in bidding for i's contract. This set may include both pure-play foundries and IDMs. $\kappa_{igt} \geq 0$ represents the shock to the buyer's utility associated with contracting with a manufacturer in the same geographic region. It can, for example, be interpreted as a government subsidy or tax exemption that benefits the buyer. The indicator function $\mathbf{1}(geog_{igt} \in geog^j_{gt})$ is equal to 1 if manufacturer j has available capacity in the same region as i, and it implies that κ_{igt} only affects the buyer's payoff from contracting with a regional manufacturer. We assume that κ_{igt} is drawn i.i.d. across i, g, t from a distribution $H(\cdot)$. Thus, a given buyer contracting with different regional manufacturers receives the same κ_{igt}. p^j_{igt} denotes the price manufacturer j charges. A transaction occurs when the j that maximizes Equation (1) results in the equation being greater than or equal to zero.

We assume that the manufacturers have a buyer-specific marginal cost of manufacturing in generation g and period t. We parametrize this cost as:

$$c^j_{igt} = \underbrace{c_{gt} + c_j - f(\boldsymbol{q^j_{gt}}, Q^j_{gt})}_{c^j_{gt}} + c^{geog}_{j,igt} - \lambda \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\}) + \omega_{j,igt}. \tag{2}$$

c_{gt} is the common cost of manufacturing generation g technology in period t, shared by all manufacturers. c_j is a manufacturer-specific component of marginal cost; this term allows for particular manufacturers to have persistent cost advantages. Q^j_{gt} captures available capacity from manufacturer j for generation g in period t. $\boldsymbol{q^j_{gt}}$ is a vector of all order quantities across buyers i to be fulfilled by firm j for generation g in period t. $f(\boldsymbol{q^j_{gt}}, Q^j_{gt})$ is a function that allows marginal costs to vary with firm capacity utilization. We constrain $\frac{\partial f(\boldsymbol{q^j_{gt}}, Q^j_{gt})}{\partial Q^j_{gt}} \geq 0$, which captures both the scale effect and the lower opportunity costs faced by manufacturers with larger capacity when accepting orders from current buyers. The first three terms do not vary across buyers within the same technology generation and time period for manufacturer j. To simplify the notation, we denote these terms by the letter c^j_{gt}.

$c^{geog}_{j,igt}$ is the region-specific manufacturing cost for buyer i's order. A manufacturer operating multiple facilities across regions may face different marginal costs in each region. λ is the re-

gional manufacturer cost advantage, which captures marginal cost savings generated by domestic manufacturing, through subsidies, lower transportation costs in the same region, and so on.[5] Note that $\lambda + \kappa_{igt}$ can be interpreted as the total tax breaks, savings on tariffs, or subsidies that are realized in a given regional buyer-seller relationship. $\omega_{j,igt} \overset{\text{iid}}{\sim} F(\cdot) = c_{gt} \cdot Gumbel(-\gamma\sigma_\omega, \sigma_\omega)$ is an idiosyncratic error term that captures the remaining variations in the marginal cost for manufacturer j to manufacture chips for buyer i. γ is the Euler–Mascheroni constant.[6]

B Negotiation Process, Information, and Timing

For a given time period t, before all negotiations begin, there is mass $I_{gt} \sim D_{gt}(\cdot)$ of buyers for generation g in period t, drawn from a demand distribution $D_{gt}(\cdot)$ that evolves over time. These buyers arrive simultaneously at t and each buyer i seeks to contract manufacturing capacity through the negotiation process with all manufacturers with available capacity. Each buyer i chooses the technology generation g and the quantity of chips q_{igt}, which is perfectly inelastic and exogenous to the bargaining process.[7] We explore allowing for elastic individual order quantities that vary with order prices in Appendix C.1.

We assume that each buyer's willingness to pay, quantity of chips ordered, and each manufacturer's available capacity and marginal cost function are common knowledge, known by all manufacturers for all buyers i and generations g in period t. Let common knowledge be represented by the letter $\boldsymbol{s}_{gt}$. *The only private information in this game is* $\omega_{j,igt}$, which is each manufacturer's own private idiosyncratic cost specific to each buyer contract. Each manufacturer only knows their own cost and the common distribution $F(\cdot)$ from which all idiosyncratic costs are drawn.

We assume the reservation price is equal to $v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\})$. Manufacturers participate in bidding for the buyer's contract only if they can provide positive utility; that

[5] If the manufacturer has available capacity in the same region as the buyer, we assume that the order will be manufactured in that region. If not, we assume that the order will be manufactured in a randomly selected region where the manufacturer has capacity for generation g in period t.

[6] We adjust by $-\gamma\sigma_\omega$ so that the expected mean (i.e., inclusive value) of the distribution of $\omega_{j,igt}$ is 0 as in Allen et al. (2019).

[7] We estimate individual order demand elasticities in Appendix C.1C and find that they are statistically indistinguishable from zero, with tight confidence intervals, consistent with perfectly inelastic demand. It is also reasonable to assume that the quantity and product characteristics demanded are inelastic because the quantity and the kind of chips a chip designer needs largely depend on downstream chip demand from electronics manufacturing. According to IC Insights (2022), the semiconductor content's value share in electronic systems was only 21-26% between 2004 and 2015, which corresponds to our sample period. Another justification for this assumption is the periodic capacity shortages and gluts observed in this industry. If quantities were elastic, manufacturers could have adjusted prices to smooth out the shortages and gluts.

is, when $v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\}) \geq c^j_{gt} + c^{geog}_{j,igt} - \lambda \cdot \mathbf{1}(geog_{igt} \in geog^j_{gt}) + \omega_{j,igt}$. If no manufacturer can offer a positive net utility to the buyer, no transaction takes place, and the buyer exits. The negotiation process and the resulting equilibrium price of the game proceed as follows.

One Manufacturer Bidding: If there is only one manufacturer with available capacity participating, we assume that bargaining takes the form of one TIOLI offer made by the manufacturer to the buyer. The transaction takes place and the manufacturer charges:

$$p^{(1)}_{igt} = v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\}). \tag{3}$$

Multiple Manufacturers Bidding: If there are multiple manufacturers with available capacity participating, the buyer runs what is strategically equivalent to a second-price procurement auction, the sealed-bid counterpart of a descending (reverse) English auction. The winning manufacturer, which we index as $j = (1)$, charges:

$$p^{(1)}_{igt} = c^{(2)}_{gt} + c^{geog}_{(2),igt} - \mathbf{1}(geog_{igt} \in \{geog^{(2)}_{gt}\})(\lambda + \kappa_{igt}) + \omega_{(2),igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog^{(1)}_{gt}\}), \tag{4}$$

where $j = (2)$ indexes the manufacturer which provides the second-highest utility for the buyer (which takes into account both marginal cost and regional match benefits κ_{igt}).

C Expected Manufacturer Profit

If there is only one manufacturer with available capacity bidding, that manufacturer makes the following profit from each unit ordered:

$$\pi_{j,igt}(\boldsymbol{s}_{gt}) = v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\}) - (c^j_{gt} + c^{geog}_{j,igt} - \lambda \cdot \mathbf{1}(geog_{igt} \in \{geog^j_{gt}\}) + \omega_{j,igt}). \tag{5}$$

If there are multiple (> 1) manufacturers with available capacity bidding, the expected profit of the winning firm $j = (1)$ per unit quantity ordered is:

$$E[\pi_{j,igt}(\boldsymbol{s}_{gt})|j = (1)] = E[p^{(1)}_{igt}(\boldsymbol{s}_{gt})] - (c^{(1)}_{gt} + c^{geog}_{(1),igt} - \lambda \cdot \mathbf{1}(geog_{igt} \in \{geog^{(1)}_{gt}\}) + \omega_{(1),igt}), \tag{6}$$

Because the idiosyncratic cost shocks are i.i.d. extreme value, each manufacturer's probability of winning, $P_{j,igt}[j = (1)|\boldsymbol{s}_{gt}]$, takes a logit form in effective relative costs. The buyer's willingness to pay drops out in that expression because it is common across all participating manufacturers.

The expected winning price is given by the expected second-lowest effective cost among bidders. Appendix B.1 reports the corresponding closed-form expressions.

The order-level expected profit for j contracting with i is the quantity q_{igt} multiplied by the expected profit upon winning, $E[\pi_{j,igt}(\boldsymbol{s}_{gt})|j=(1)]$, and the probability of winning (conditional on j participating in bidding for i's contract). The total expected profit of manufacturer j in period t sums the order-level expected profits across buyers where j participates in bidding and across technology generations and subtracts other fixed cost of goods sold, such as labor costs and manufacturing overhead, which we represent by C_{jt}:

$$E[\pi_{j,t}(\boldsymbol{s}_t)] = -C_{jt} + \sum_g \sum_i E[\pi_{j,igt}(\boldsymbol{s}_{gt})|j=(1)] \cdot P_{j,igt}[j=(1)|\boldsymbol{s}_{gt}] \cdot q_{igt}. \tag{7}$$

5 Model Estimation

5.1 Parameterization, Identification, and Calibration

This subsection specifies a functional form for each object of the model, states the variation that identifies each parameter, and notes the quantities we calibrate. We proceed in the order of the profit function: the marginal cost components of Equation (2), then fixed costs, then demand, and finally the capacity available for contract manufacturing. Table 4 summarizes the resulting mapping from moments to parameters.

Common marginal costs At any given t, as g becomes more cutting-edge, the common cost of manufacturing c_{gt} becomes larger. At any given g, as t goes on, c_{gt} becomes cheaper. Thus, we parameterize c_{gt} as a function of generation g and period t:

$$c_{gt} = \alpha_0 g^{\alpha_g} t^{\alpha_t}. \tag{8}$$

The parameters pinning down c_{gt} are identified by the variation in average prices by generation and over time. In particular, α_g captures how relative to some base cost α_0 corresponding to our earliest observed generation and time period, average costs change across generations, and α_t similarly captures how average costs vary across time.[8]

Identifying firm versus place-based cost advantages c_j is parameterized as a fraction of c_{gt}, that is, $c_j = \tilde{c}_j \cdot c_{gt}$. A cost-competitive manufacturer may have a negative $\tilde{c}_j$, while a high-cost manufacturer may have a positive $\tilde{c}_j$. The value of $\tilde{c}_j$ for each j is identified through

[8] $g = 5$ and $t = 2003Q3$ are normalized to a value of zero in this equation.

variation in the manufacturer's revenue share. We assume that smaller pure-play contract manufacturers for which we do not observe revenue (they collectively account for less than 10% of the contract manufacturing market) share the same value of $\tilde{c}_j$, which we normalize to zero. Additionally, we assume that IDMs share the same, relatively large positive value for $\tilde{c}_{IDM}$, given that their fabs are not optimized for handling the diverse product mix characteristic of contract manufacturing.[9] The revenue share of IDMs within the total contract manufacturing market identifies $\tilde{c}_{IDM}$.

We parametrize $c^{geog}_{j,igt}$ to take on four possible values: c^{CN}, c^{NA}, c^{ROW}, and c^{TW}, representing the region-specific manufacturing cost for mainland China, North America, Rest of the World, and Taiwan respectively. Each of these four terms is parameterized as a fraction of c_{gt}, that is $c^{geog} = \tilde{c}^{geog} \cdot c_{gt}$, and $\tilde{c}^{ROW}$ is normalized to zero.

Separating the location components $\tilde{c}^{geog}$ from the firm components $\tilde{c}_j$ is central to our counterfactuals because it determines how much of TSMC's efficiency can be carried to a new location: relocating a fab preserves the firm component but replaces the location component with that of the destination. Two sources of variation separate these components, and a third pins down the scale on which both are measured.

First, firm identity is observed in the revenue and margin data. Differences in these outcomes across manufacturers producing in the same region therefore help identify the firm components. Second, many manufacturers operated facilities in multiple regions during our sample period, providing within-firm variation in exposure to different location components. For example, during our sample period, TSMC operated fabs in Taiwan, mainland China, and North America, while UMC operated fabs in Taiwan, mainland China, Singapore, and Japan. As a manufacturer's capacity shifts across regions, the cost efficiency of its overall capacity portfolio changes even though its firm component remains fixed. The resulting changes in its cost competitiveness generate corresponding changes in its revenue relative to that of its competitors. This helps identify the location components separately from the firm components. Variation in where each manufacturer produces thus helps distinguish the cost advantage associated with being TSMC or UMC from the cost advantage associated with producing in Taiwan rather than elsewhere.

[9] The only exception is Samsung, which had dedicated foundries for contract manufacturing during our sample period. We estimate its $\tilde{c}_{samsung}$ separately from the rest of the IDMs.

Third, the price moments, formed by region within each technology generation and period and used to pin down the common cost c_{gt}, help fix the scale of $\tilde{c}_j$ and $\tilde{c}^{geog}$. Because both the firm and location components are parameterized as fractions of the common cost, the level of c_{gt} determines what a given difference in revenues implies for $\tilde{c}_j$ and $\tilde{c}^{geog}$. Prices are informative about this common cost level rather than providing a direct measure of the winning manufacturer's cost, because the transaction price in our model equals the second-lowest effective cost among bidders rather than the cost of the manufacturer that fulfills the order. All parameters are estimated jointly, so these sources of variation operate simultaneously rather than sequentially.

Scale effects and opportunity cost For computational tractability, we assume that $f(\boldsymbol{q_{gt}^j}, Q_{gt}^j) = \zeta \cdot ln(Q_{gt}^j)$, where $\zeta \geq 0$. This specification allows manufacturers with larger installed capacities to be more cost efficient, reflecting both economies of scale and the lower opportunity cost of accepting orders from current buyers. The parameter ζ captures the decline in marginal cost as a manufacturer's installed capacity for a technology generation increases, and we parameterize it as a fraction of c_{gt}, that is, $\zeta = \tilde{\zeta} \cdot c_{gt}$. Variation in utilization across manufacturers and over time, driven by changes in installed capacity, identifies $\tilde{\zeta}$.

Under this assumption, the model does not impose a hard capacity constraint. Instead, marginal cost declines as installed capacity increases, which induces manufacturers with larger capacity to win more orders and fill that capacity. A limitation of this approach is that a manufacturer could, in principle, receive orders far exceeding its installed capacity, whereas in the data, utilization never exceeds 110%. Addressing this fully would require solving for the equilibrium price at which the equilibrium quantity is consistent with installed capacity. However, characterizing equilibrium bidding strategies when marginal cost depends on quantity and utilization, rather than on installed capacity alone, is non-trivial both theoretically and computationally. We therefore retain this simplification in the baseline model and explore relaxing it in Appendix C.2, where we show that our main policy counterfactual results are robust to imposing a capacity cap corresponding to 110% utilization.

Proximity benefits We assume that $\kappa_{igt} = \tilde{\kappa}_{igt} \cdot c_{gt}$, where $\tilde{\kappa}_{igt}$ is drawn from an exponential distribution with rate parameter σ_κ (so that $E[\tilde{\kappa}_{igt}] = 1/\sigma_\kappa$), and $\lambda = \tilde{\lambda} \cdot c_{gt}$, where $\tilde{\lambda}$ is a constant parameter. Together, these two terms capture the benefits of geographic proximity

between a manufacturer and its buyer, which industry participants attribute to lower shipping and logistics costs, easier coordination between design and manufacturing teams, and better protection of intellectual property. The two operate on different sides of the transaction: λ is a cost saving to the manufacturer from serving a same-region buyer and is common across matches, while κ_{igt} is extra value to the buyer from contracting with a same-region manufacturer and varies across matches, reflecting heterogeneity in how much individual buyers value co-location. In equilibrium the two enter bids symmetrically, so a same-region manufacturer enjoys an effective advantage of $\lambda + \kappa_{igt}$ over otherwise identical rivals in the auction for buyer i's contract. $\tilde{\lambda}$ is identified by differences in the revenue share from Chinese buyers between mainland Chinese and Taiwanese manufacturers, while σ_κ is separately identified by variation in this share across periods and manufacturers: a common cost saving shifts the average share, whereas dispersion in the share across otherwise similar manufacturers and periods reflects the distribution of buyer-side match values.

Idiosyncratic buyer-manufacturer-specific cost $\omega_{j,igt}$ captures variation in prices not explained by generation-period, manufacturer-specific, or location-specific components. The remaining variation in prices identifies σ_ω.

Fixed costs The other fixed cost of goods sold, C_{jt}, is parameterized as follows:

$$C_{jt} = \sum_{k \in K_{jt}} C^{geog} \cdot ln(Q^k_{jt}), \tag{9}$$

where Q^k_{jt} is the installed capacity of facility k that manufacturer j operates. C^{geog} varies by region, with three region-specific values: C^{CN} (China), C^{NAROW} (North America and ROW), and C^{TW} (Taiwan). The C^{geog} terms are identified through occurrences of negative gross margins among manufacturers with facilities in different regions. Because we do not observe margins for any manufacturer whose installed capacity is predominantly located in North America, we combine North America and ROW into a single category. The total fixed cost is the sum of these facility-specific costs.

Demand To simplify computation, we assume that buyer willingness to pay (v_{igt}) is high enough that the manufacturer participation constraint never binds, so all manufacturers with available capacity for a technology generation and period participate in bidding. A consequence

is that we cannot identify buyer willingness to pay itself.[10] We discuss the implications of this assumption in Section 8.

The demand distribution $D_{gt}(\cdot)$ is parameterized as a deterministic categorical distribution $d_{gt}(geog_i|\cdot)$ of contract share by buyer geographical location multiplied by a fixed distribution of contract share by distance to technology frontier $d(g_t^{max} - g_{it})$, scaled by the WSTS semiconductor end product sales time series and a demand scaling parameter σ_D: $D_{gt}(\cdot) = \sigma_D \cdot WSTS_{t+1} \cdot d(g_t^{max} - g_{it}) \cdot d_{gt}(geog_i|\cdot)$.[11] Because we observe both price and quantity for a representative set of orders, total revenue in the contract manufacturing market helps identify the demand scaling parameter σ_D.

We assume that the share of contracts by distance to technology frontier $d(g_t^{max} - g_{it})$ does not change across time and estimate these shares by aggregating orders by distance to frontier in our entire orders data.

For simplicity, we assume buyer locations fall into three categories: mainland China (CN), North America (NA), and the rest of the world (ROW), since mainland China and North America are at the center of the recent policy changes. The share of contracts from each buyer region, $d_{gt}(geog_i|\cdot)$, follows a multinomial logit in the order's distance to the technology frontier and in time. For example, the share of contracts from Chinese buyers is parameterized as

$$d_{gt}(geog_i = CN|g_{it}, t) = \frac{exp(l_0^{CN} + l_1^{CN}(g_t^{max} - g_{it}) + l_2^{CN}t)}{1 + \sum_{geog}^{\{CN,NA\}} exp(l_0^{geog} + l_1^{geog}(g_t^{max} - g_{it}) + l_2^{geog}t)} \tag{10}$$

The shares of North American and ROW buyers are defined analogously, with ROW serving as the reference category; Appendix B.2 reports the full functional forms. These functional forms are informed by industry knowledge: North American buyers on average demand more cutting-edge technologies, and the share of Chinese buyers has steadily risen while the share of North American buyers has fallen since the 2000s. For each region, the logit intercept is identified by the average revenue share from that region's buyers across manufacturers, the

[10] In principle, v_{igt} could be identified from monopoly situations, where the equilibrium price equals the buyer's willingness to pay plus any regional manufacturer benefit (κ_{igt}), or from variation in the set of potential bidders over time, as in Kong (2022). We observe no monopoly generation-periods in our data, and the second approach would require solving the model for every possible set of entrants at each candidate parameter vector, which is computationally very costly in our setting, since the set of potential entrants includes many small firms beyond the industry leaders we focus on.

[11] We use a one-quarter lead of the WSTS series due to a typical lag of one quarter between manufacturing demand and sales data. Manufacturers often see demand shifts from chip designers a quarter before these are reflected in chip designers' sales. Notably, manufacturing orders usually peak in the second quarter and dip in the third, while semiconductor sales peak in the third quarter and fall in the fourth.

distance coefficient by comparing these shares across manufacturers with different technology portfolios (e.g., TSMC vs. UMC, SMIC vs. Hua Hong), and the time trend by tracking the shares for the same manufacturers over time.

Capacity available for contract manufacturing For a pure-play foundry, all of its capacity is available for contract manufacturing. For an IDM, we assume that a random proportion m_{jt}^k of each facility's capacity is available for contract manufacturing in period t, drawn each period from a distribution centered at zero, since IDM capacity is, in expectation, fully utilized by the IDM's own products. We calibrate the standard deviation of the underlying draw at $\sigma_{\tilde{m}} = 0.05$, as fabs usually need utilization above 95% to justify their fixed cost investment; $\tilde{m}$ is calibrated rather than estimated because it cannot be separately identified from the IDM-specific marginal cost $\tilde{c}_{IDM}$. Appendix B.2 gives the exact specification.

5.2 Estimation Method and Moments

Let $\boldsymbol{\theta} = \{\alpha_0, \alpha_g, \alpha_t, \{\tilde{c}_j\}, \{\tilde{c}^{geog}\}, \tilde{\zeta}, \tilde{\lambda}, \sigma_\omega, \{C^{geog}\}, \sigma_\kappa, l_0^{CN}, l_1^{CN}, l_2^{CN}, l_0^{NA}, l_1^{NA}, l_2^{NA}, \sigma_D\}$ represent the parameters to be identified. We estimate these parameters using the method of simulated moments (McFadden, 1989), which searches the parameter space to minimize the differences between simulated and actual data patterns. Specifically, at each iteration of the estimation algorithm, we use the current combination of $\boldsymbol{\theta}$ parameters to first simulate the number of buyers, $I_{gt} \sim D_{gt}(\cdot)$, based on our parameterization. Next, we simulate a series of multilateral negotiations between buyers and the available manufacturers using our negotiation game. These negotiation outcomes generate the simulated data moments, which we compare to the actual data moments. The parameter combination that minimizes these differences provides the recovered model-implied estimates for $\boldsymbol{\theta}$. The data moments matched by our estimation routine are shown in Table 4.

To estimate the optimal model parameters, we minimize the following weighted distance function between observed and simulated moments:

$$L(\boldsymbol{\theta}) = \left[\frac{m - \hat{m}(\boldsymbol{\theta})}{\bar{m}}\right]' W \left[\frac{m - \hat{m}(\boldsymbol{\theta})}{\bar{m}}\right], \tag{11}$$

where $L(\boldsymbol{\theta})$ is the objective function to be minimized. m is the vector of empirical moments calculated from observed data. $\hat{m}(\boldsymbol{\theta})$ is the vector of simulated moments generated by the model using parameters $\boldsymbol{\theta}$.

Table 4: Moment Conditions

Moment	Description	Parameters Identified	Data Source
$margin_{jt}(\boldsymbol{\theta})$	The gross margin of manufacturer j in period t, for the subset of firms with available margin data.	$\{C_{geog}\}$	Quarterly reports
$pmean_{geog,gt}(\boldsymbol{\theta})$	The average price for each technology generation, period, and fab region.	α_0, α_g, α_t	GSA orders
$pvar_{geog,gt}(\boldsymbol{\theta})$	The variance of prices within each technology generation, period, and fab region.	σ_ω	GSA orders
$revenue_{jy}(\boldsymbol{\theta})$	Annual revenue for manufacturer j, available for a group of leading contract manufacturers, including both pure-play foundries and IDMs.	$\{\tilde{c}_j\}$, $\{\tilde{c}^{geog}\}$	IC Insights
$revenue_{jt}(\boldsymbol{\theta})$	Quarterly revenue for manufacturer j, available for a subset of pure-play foundries.	$\{\tilde{c}_j\}$, $\{\tilde{c}^{geog}\}$	Quarterly reports
$revenue_{y}(\boldsymbol{\theta})$	Annual contract manufacturing market revenue, available for a subset of years.	σ_D	IC Insights
$revshare_{IDM,y}(\boldsymbol{\theta})$	The revenue share coming from IDMs among all contract manufacturers, available for a subset of years.	$\tilde{c}_{IDM}$	IC Insights
$utilization_{jt}(\boldsymbol{\theta})$	The average capacity utilization rate for manufacturer j, available for a subset of firm-quarters.	ζ	Quarterly reports
$revshare_{jt,CN}(\boldsymbol{\theta})$	Revenue share from Chinese buyers for manufacturer j, available for a subset of firm-quarters.	σ_κ, $\tilde{\lambda}$, l_0^{CN}, l_1^{CN}, l_2^{CN}	Quarterly reports
$revshare_{jt,NA}(\boldsymbol{\theta})$	Revenue share from North American buyers for manufacturer j, available for a subset of firm-quarters.	l_0^{NA}, l_1^{NA}, l_2^{NA}	Quarterly reports

Note: Every simulated moment additionally takes the SEMI fab capacity data as an input, since the set of manufacturers and their regional capacity composition determine the bidder sets in each simulated market; the revenue moments therefore embed the within-firm regional variation that separates $\{\tilde{c}_j\}$ from $\{\tilde{c}^{geog}\}$. The demand distribution is scaled by the WSTS semiconductor sales series. For information about data availability, please see Appendix D.

$\bar{m}$ is a vector of scaling factors used to normalize the moments. The moments in our data have widely different scales. For instance, gross margins and utilization rates are small numbers (e.g., 0.3 or 0.98), while revenue figures are in millions of dollars. To make these moments comparable, we normalize each moment by dividing it by the mean value of that moment type in the observed data. For example, for manufacturer j in period t, the margin moment is $\frac{\text{observed margin}_{jt} - \text{simulated margin}_{jt}}{\text{mean observed margin}}$. This normalization is captured by the $\bar{m}$ term in the objective function.

The weighting matrix W balances the contribution of different moment groups in the objective function. Since we have unequal numbers of observations across moment types, we use weights to ensure each group contributes equally to the estimation. The matrix W is diagonal, with entries being the inverse of the number of moments in each group. For example,

manufacturer-year revenue has 139 observations in our sample period, so these moments receive a weight of 1/139. IDM revenue shares have 8 observations, so these moments receive weight 1/8. This weighting scheme prevents the objective function from being dominated by moment groups that simply have more observations, ensuring all moment conditions contribute meaningfully to parameter estimation. Consistency of our estimator does not depend on this choice of weighting matrix: method of simulated moments estimators are consistent for any positive definite weighting matrix (Hansen, 1982; McFadden, 1989; Pakes and Pollard, 1989).

6 Estimation Results

We present our estimated parameters in Table 5. The first three parameters capture the evolution of the common, generation-specific marginal manufacturing cost over time. As expected, the estimate for α_g is greater than 0, indicating that more advanced technology is more expensive to manufacture at any given time, while the estimate for α_t is smaller than 0, suggesting that manufacturing costs decline as the technology matures. The estimates imply that for $g = 5$ and $t = 2003Q3$, which are normalized to zero in Equation 8, the common marginal manufacturing cost is \$2811.05 per m^2.

Firm-specific and region-specific marginal costs capture deviations from the common marginal cost. The baseline category omitted in the estimation consists of smaller pure-play foundries located in the Rest of the World (i.e., outside mainland China, Taiwan, or North America). For these firms, the evolution of marginal cost is reflected by the common marginal cost estimates. For other firms, including larger pure-play foundries and IDMs, the firm- and region-specific estimates capture deviations from this common component.

For example, $\tilde{c}_{TSMC} = -0.1396$, $\tilde{c}_{TW} = -0.0940$, and $\tilde{c}_{NA} = 0.0003$. This implies that, on average, the marginal cost of a TSMC fab located in Taiwan is lower than that of a small pure-play foundry located in North America by $[(-0.1396 - 0.0940) - 0.0003] \times 100\% = -23.39\%$ of the common marginal cost for that technology generation and quarter, all else equal. Similarly, the marginal cost of a TSMC fab located in Taiwan is lower than that of a TSMC fab located in North America by $[(-0.0940) - 0.0003] \times 100\% = -9.43\%$ of the common marginal cost for that technology generation and quarter, all else equal.

σ_κ and λ capture the savings and benefits of contracting with a manufacturer in the same region as the buyer. The estimates imply that this benefit on average is $(1/228528.0892 +$

0.0796) $\times$ 100% = 7.96% of the common marginal cost under the relatively stable policy environment between 2004 and 2015.

Estimates for region-specific fixed cost of manufacturing captures how much this cost increases as fabs located in the respective regions expand capacity. For a 1% increase of capacity in a fab located in mainland China, the estimate implies that the fixed cost of manufacturing would increase by 1% $\times$ 155030.0707 = \$1550.30 per quarter, all else being equal.

Table 5: Estimation results

Parameter Description	Parameter	Coefficient	90% confidence interval
Common marginal cost	α_0	2811.0498	[2725.6116, 2910.9951]
	α_g	0.4619	[0.4462, 0.4711]
	α_t	−0.3209	[-0.3294, -0.3128]
Firm-specific marginal cost	$\tilde{c}_{tsmc}$	−0.1396	[-0.1572, -0.1207]
	$\tilde{c}_{umc}$	0.0291	[0.0157, 0.0401]
	$\tilde{c}_{globalfoundries}$	0.0065	[-0.0630, 0.1594]
	$\tilde{c}_{samsung}$	−0.0036	[-0.0727, 0.0891]
	$\tilde{c}_{smic}$	0.0126	[-0.0107, 0.0497]
	$\tilde{c}_{chartered}$	−0.1761	[-0.2269, -0.1274]
	$\tilde{c}_{huahong}$	0.0666	[0.0402, 0.0988]
	$\tilde{c}_{tower}$	0.0178	[0.0047, 0.0340]
	$\tilde{c}_{powerchip}$	0.0537	[0.0399, 0.0701]
	$\tilde{c}_{vanguard}$	−0.0239	[-0.0445, -0.0041]
	$\tilde{c}_{IDM}$	0.0408	[0.0290, 0.0544]
Region-specific marginal cost	$\tilde{c}_{CN}$	−0.0129	[-0.0186, -0.0084]
	$\tilde{c}_{NA}$	0.0003	[0.0000, 0.0007]
	$\tilde{c}_{TW}$	−0.0940	[-0.1030, -0.0875]
Regional manufacturing savings/benefits	σ_κ	228 528.0892	[4356.5500, 829780.8450]
	$\tilde{\lambda}$	0.0796	[0.0780, 0.0800]
Region-specific fixed cost	C_{CN}	155 030.0707	[38168.4502, 256013.4074]
	C_{NAROW}	876 742.8324	[744002.6066, 1003001.0859]
	C_{TW}	424 324.7158	[301652.2431, 520122.9124]
Buyer location parameters	l_0^{CN}	0.3720	[0.1997, 0.6363]
	l_1^{CN}	−0.2028	[-0.2194, -0.1868]
	l_2^{CN}	0.0683	[0.0382, 0.0869]
	l_0^{NA}	9.5289	[9.2160, 9.9743]
	l_1^{NA}	−1.7416	[-1.7866, -1.6895]
	l_2^{NA}	−0.1149	[-0.2594, -0.0138]
Effect of scale on marginal cost	$\tilde{\zeta}$	0.0374	[0.0359, 0.0395]
Idiosyncratic variation in marginal cost	σ_ω	0.1525	[0.1425, 0.1584]
Demand scaling parameter	σ_D	27.3083	[25.8972, 28.7691]

Note: We simulate and estimate the parameters 50 times, where the number of buyers, $\omega_{j,igt}$, κ_{igt}, and other variables in the simulated data are randomly drawn for each simulation. The reported coefficients are mean values across the 50 estimations. The 90% confidence interval is constructed by taking the 5th percentile and 95th percentile of the estimates across the 50 estimations. The confidence interval accounts for simulation errors. Since we only have one global market for contract manufacturing, we do not draw with replacement from the observed data; thus, these confidence intervals do not account for sampling errors in the observed data.

The buyer location parameters capture how the share of buyers from mainland China, North America, and the Rest of the World (including Taiwan) changes with technology generation and

time. Three key patterns emerge from these estimates: (1) North American buyers represent a large share of the buyer pool, as indicated by the large positive estimate of l_0^{NA} and the relatively small estimate of l_0^{CN}. (2) More advanced technology generations have a higher proportion of North American buyers, reflected by the large negative estimate of l_1^{NA} and the relatively small estimate of l_1^{CN}. (3) The share of mainland Chinese buyers has been increasing over time at the expense of North American buyers, as shown by the positive estimate of l_2^{CN} and the negative estimate of l_2^{NA}. These patterns are consistent with industry knowledge.

The scale effect $\tilde{\zeta}$ is estimated at 0.0374, implying that a 1% increase in a fab's installed capacity reduces the marginal manufacturing cost by $1\% \times 0.0374 = 0.0374\%$ of the common marginal cost, all else being equal. σ_ω is estimated at 0.1525, indicating a substantive buyer-manufacturer specific idiosyncratic marginal cost. One standard deviation of this idiosyncratic cost is $\sqrt{\pi^2/6 \times 0.1525^2} \times 100\% = 19.56\%$ of the common marginal cost.

7 Policy Counterfactuals: TSMC Fab Location Case Study

We study a detailed case of TSMC Fab 14's location choice amid various potential policy combinations. In this exercise, we perturb the location of a single facility, holding fixed all of TSMC's other investment decisions and the equilibrium industry structure described in Section 2.3. The location comparison can thus be read as a partial derivative of the firm's investment problem with respect to the location choice, and the policy scenarios as cross-partial effects that measure how each policy alters that comparison. We find that relocating to the U.S. generates billions in flow static profit losses for TSMC under different combinations of domestic and foreign policies, with the exception of import tariffs. We also find that fab losses from locating in the U.S. increase as technologies mature. This suggests that policymakers would need additional subsidies or other policies to incentivize TSMC to continue its U.S. operations in the medium to long run after initial investments are sunk.

To calculate our counterfactuals, we simulate TSMC's profit streams $E[\pi_{jt}(\boldsymbol{s}_t;\boldsymbol{\theta})]$ for $t = 2004Q2, ..., 2015Q3$ (from the quarter Fab 14 began production through the end of our data period) under these policy scenarios, comparing outcomes for Taiwan versus U.S. locations. We also employ a machine learning approach to simulate the additional investment costs of building and maintaining U.S. capacity if Fab 14 were located in the U.S. These costs include both the initial construction and equipment expenses, beginning with the start of Fab 14's construction

in 2000, as well as the subsequent costs of expanding and upgrading the facility through the end of our sample period in 2015. Using the fab investment data from Table 3, we train a multi-layer perceptron neural network with fab characteristics (e.g., manufacturer identity, location, capacity, and technology) as input features and investment amount as the target variable. This model achieves superior in-sample and out-of-sample performance in predicting investment costs based on the characteristics of the fab investment project (see Appendix E for details). Finally, we then compare profit changes against these estimated investment costs while accounting for potential subsidies to determine the optimal location under each policy scenario.

A Baseline Scenario

In the baseline simulation, we simulate TSMC's profit from 2004 to 2015 using cost and demand parameters estimated from observed data, $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, while keeping Fab 14's location fixed at its actual site in Taiwan. This approach enables us to evaluate our model's fit by comparing the simulated profit totals in this baseline scenario with the actual profit totals.

Table 6: TSMC Fab Investment and Profit: Baseline vs. U.S. Relocation

	Actual		Simulated		U.S. subsidy	
	Investment	Profit	Investment	Profit	Direct	ITC
	(1)	(2)	(3)	(4)	(5)	(6)
A. Fab 14	−\$6.6B	\$77.1B	−\$6.7B	\$74.0B	NA	NA
in Taiwan (baseline)			[5.8,7.5]	[70.3,78.6]		
B. Fab 14	NA	NA	−\$7.8B	\$72.1B	\$0.8B	\$2.0B
relocated to the U.S.			[6.7,8.8]	[68.5,76.9]		
Difference (B−A)			−\$1.2B	−\$1.8B	+\$0.8B	+\$2.0B
			[−1.8,−0.6]	[−2.5,−1.1]		

Notes: Actual and simulated investment are undiscounted sums of investments in Fab 14 from 2000 to 2015. Actual and simulated profit are undiscounted sums of profits for all of TSMC from 2004Q2 to 2015Q3; we do not isolate Fab 14's profitability due to potential substitution effects across production facilities. Simulated profits are means across the 50 sets of $\widehat{\boldsymbol{\theta}}$ estimates described in the notes of Table 5, and simulated investments are mean predictions from 30 independently trained MLP neural network models using different random seeds; the 90% confidence intervals in square brackets are based on the 5th and 95th percentiles of the respective simulations. Columns (5) and (6) report the direct subsidy and the Investment Tax Credit (ITC), the primary manufacturing incentives under the CHIPS and Science Act, both treated as lump sums at face value: the direct subsidy is set at 10% of total investment, matching TSMC Arizona's \$6.6 billion award on a \$65 billion investment, and the ITC at the 25% statutory rate under which that award was made. Appendix A.2 details the statutory basis of the ITC and the tax offsets from which these face values abstract.

The first row of Table 6 presents TSMC Fab 14's actual total investment from 2000 to 2015, alongside TSMC's actual profits from 2004Q2 (when Fab 14 began production) through 2015Q3 (the end of our sample period). The same row also shows the simulated investment and

profit for these periods, generated using our parameter estimates. The results demonstrate a strong fit between the model and the data. The simulated investment amount—predicted based on Fab 14's project characteristics, such as capacity expansion and technology generation—is nearly identical to the actual total investment. The simulated total profit for TSMC during 2004Q2 to 2015Q3 is \$74.0 billion, with a 90% confidence interval of $[70.3, 78.6]$. The actual profit of \$77.1 billion falls within this interval. Overall, the simulated investment and profit closely match the actual figures. We use these simulated results as our baseline for comparison with counterfactual simulations under alternative policies.

B Fab 14 Relocated to the U.S. Under Capacity Subsidies

In this simulation, we relocate Fab 14 to the U.S. in response to capacity subsidies similar to those offered under the CHIPS and Science Act. These U.S. subsidies fall into two categories: a direct subsidy and an investment tax credit (ITC). The ITC, established under Section 48D of the Internal Revenue Code, entitles the firm to a credit equal to 25% of its qualified investment, which includes the buildings and equipment integral to the fab, and offsets the firm's tax liability dollar for dollar. Because the statute's elective payment provision makes the credit refundable even when the firm owes little U.S. tax, the ITC functions as a cash payment proportional to investment, much like the direct subsidy. We therefore treat both as lump-sum payments received at face value; neither affects the marginal cost of production or the marginal cost and benefit of operating in a specific geography. To isolate the effect of Fab 14's location on market competition, we hold constant all other dimensions of the investment projects and competitive environment. The only change is Fab 14's location. Accordingly, as in the baseline simulation, we simulate TSMC's profit from 2004 to 2015 using cost and demand parameters estimated from observed data, $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. The second row of Table 6 presents the simulation results. First, we find that the investment cost of building, expanding, and upgrading Fab 14 in the U.S. exceeds that of equivalent projects in Taiwan by \$1.2 billion. In addition, U.S.-based manufacturing is less cost-effective, resulting in an estimated profit loss of \$1.8 billion over the period, 83% which is driven by our estimated differences in marginal costs.[12]

[12]We can decompose profit loss into variable profit and fixed cost components. Losses from lower variable profits in the U.S. location make up for 83% of total losses, and the remaining 17% of losses are due to higher U.S. fixed costs (COGS).

C Fab 14 in Taiwan vs. the U.S. Under the Current Policy Regime

Our main policy exercise evaluates Fab 14's location under the combination of policies in effect during our writing. To our knowledge, the primary policies altering the competitive landscape of contract semiconductor manufacturing as of the end of 2024 were U.S. export controls on sales to Chinese buyers and Chinese tax exemptions for domestic manufacturers; we refer to their combination as the current policy regime. We acknowledge that the current policy environment is prone to rapid changes. To interpret the combined effects, we also simulate each policy separately. Columns (1) and (2) of Table 7 report these single-policy scenarios, and column (3) reports the current regime.

We model the export controls after Industry and Security Bureau (2022), which restricted the export of chips produced using 16/14-nanometer or more advanced technology nodes to Chinese buyers. TSMC is affected by such controls due to its reliance on U.S.-origin technologies and equipment in its manufacturing processes. The rule took effect when TSMC's most advanced technology in volume production was 3nm, representing a four-generation gap, so in our simulations we prohibit TSMC from contracting with mainland Chinese buyers for technology generations within four generations of its frontier technology in each period.

Table 7: Simulated TSMC Profits in Taiwan vs. the U.S. Under Various Policy Regimes

	U.S. Export Controls	CN Tax Exemptions	U.S. Export Controls CN Tax Exemptions	U.S. Import Tariffs U.S. Export Controls CN Tax Exemptions
	(1)	(2)	(3)	(4)
Fab 14	\$67.8B	\$70.6B	\$65.0B	\$65.0B
in Taiwan	[64.6,72.2]	[66.6,75.5]	[61.4,69.6]	[61.4,69.8]
Fab 14	\$66.3B	\$68.8B	\$63.4B	\$70.9B
relocated to the U.S.	[63.1,70.8]	[64.8,73.5]	[59.9,67.9]	[67.2,75.8]
Difference	−\$1.6B	−\$1.8B	−\$1.6B	+\$5.9B
	[−2.3,−0.9]	[−2.5,−1.1]	[−2.3,−0.9]	[5.0,6.7]

Notes: Column (1) simulates the 2022 U.S. export controls (Industry and Security Bureau, 2022) by prohibiting TSMC from contracting with mainland Chinese buyers for technology generations within four generations of its most advanced technology in each period. Column (2) simulates Chinese corporate tax exemptions (Ministry of Finance of the People's Republic of China et al., 2020) by reducing the region-specific marginal cost in China (c_{CN}) by 5% of the common marginal cost (c_{gt}). Column (3) combines the two policies. Column (4) adds U.S. import tariffs (York and Durante, 2025), simulated by reducing U.S. buyers' willingness to pay (v_{igt}) by 10% of the common marginal cost for fabs located outside the U.S. Appendix A.2 details the calibration of each policy. The remaining notes from Table 6 apply.

The Chinese tax exemptions described in Ministry of Finance of the People's Republic of China et al. (2020) reduce the standard 25% corporate income tax rate for chipmakers located in mainland China to between 0% and 12.5%, depending on a firm's technology portfolio and the

number of years since the policy announcement. In our model, these exemptions act as marginal cost subsidies for fabs operating in mainland China, since profit margins are determined by the gap between the lowest and second-lowest marginal costs. Because it is unclear how much of the tax savings are passed through to buyers in the form of lower bid prices, we apply a conservative calibration, reducing the region-specific marginal cost in China (c_{CN}) by 5% of the technology-generation- and time-specific common marginal cost (c_{gt}). All other cost and demand parameters are held at their estimated values.

Under the current regime, TSMC earns $65.0 billion when Fab 14 is located in Taiwan and $63.4 billion when it is relocated to the U.S. Relative to the no-policy baseline, the regime reduces TSMC's profit by $9.0 billion[13] when Fab 14 remains in Taiwan, and relocation adds a further loss of $1.6 billion.

Columns (1) and (2) decompose this decline. Export controls alone reduce TSMC's profit in Taiwan by $6.1 billion,[14] and Chinese tax exemptions alone reduce it by $3.4 billion.[15] The combined loss is slightly smaller than the sum of the two: some of the contracts lost under the combined regime, primarily those of high-end Chinese customers affected by export restrictions who could alternatively source from subsidized domestic manufacturers, would have been lost to either policy alone. Across all three scenarios, the additional loss from relocating Fab 14 is essentially unchanged, between $1.6 and $1.8 billion, indicating that the effects of these policies and of relocation are largely additive: the buyers TSMC loses to export controls would have contracted with it from either location, and the contracts lost to subsidized Chinese competition are similar regardless of Fab 14's location.

Figure 3A suggests a mechanism behind this loss. It plots the annual additional profit loss from relocating Fab 14 to the U.S. under the current regime. The reduction is uneven over time, with significantly greater losses in the later years. This pattern reflects the changing competitive environment and buyer demographics as Fab 14's technology matures. Between 2005 and 2010, the reduction in profitability was relatively minor. During this period, Fab 14 remained at the cutting edge through successive upgrades—from 130nm in 2004 to 40nm in 2008. At that stage, few manufacturers could produce similarly advanced chips, and most

[13] $74.0B in Table 6, column 4, minus $65.0B in Table 7, column 3. 90% confidence interval: [7.8, 10.2].

[14] $74.0B in Table 6, column 4, minus $67.8B in Table 7, column 1; the difference is computed from unrounded simulation values. 90% confidence interval: [5.5, 6.8].

[15] $74.0B in Table 6, column 4, minus $70.6B in Table 7, column 2. 90% confidence interval: [2.4, 4.3].

buyers of advanced nodes were located in the U.S. Although manufacturing in the U.S. entails higher marginal costs, some of these were offset by cost savings associated with being in the same region as the buyers. According to our estimates and industry knowledge, U.S.-based buyers dominated demand for these advanced technology generations. The increased marginal cost of producing in the U.S. (9.43% of the technology- and time-specific marginal cost c_{gt}) was nearly offset by cost savings associated with being in the same region as the buyers (7.96% of c_{gt}), allowing TSMC to remain competitive and earn similar margins as it would with a non-U.S. fab, resulting in only a small profit loss.

However, TSMC ceased upgrading Fab 14 after 2008. As the fab's technology aged and fell further behind the frontier, competition intensified, with more manufacturers capable of producing these mature technologies and a growing share of buyers coming from outside the U.S. This is precisely the maturation pattern documented in Section 2.3: entry into Fab 14's generations accumulated as the frontier advanced past them, and the demand that remained shifted toward Asian buyers. In this context, a U.S.-based Fab 14 became increasingly uncompetitive, leading to significant profit losses in the later years of operation.

These findings suggest that shifts in the competitive environment and buyer demographics can significantly affect the profitability of U.S.-based fabs as their technology matures. If U.S. policymakers seek to incentivize TSMC to maintain its typical fab strategy—building cutting-edge facilities, upgrading them several times, and then allowing them to mature—they may need to consider providing support at later stages of the fab's lifecycle.

Incorporating the up-front investment subsidies reported in Table 6, a 10% direct investment subsidy (roughly comparable to the support received by TSMC's Arizona fabs, \$6.6 billion out of \$65 billion in total investment, or 10.2%) together with a 25% ITC amounts to \$2.7 billion in total incentives for locating the fab in the U.S. These subsidies more than offset the \$1.2 billion increase in investment costs and come close to covering the combined \$2.8–\$3.0 billion in additional investment costs and profit losses within our sample window under the policy regimes in Columns (1)–(3) of Table 7. Their timing, however, differs sharply from that of the losses: the subsidies are provided up front, whereas the profit losses are concentrated later in the facility's life and continue to accrue beyond the end of our sample period.

Viewed through the initial participation decision alone, up-front subsidies of this magnitude

can therefore rationalize the decision to build, consistent with TSMC's decision to construct the Arizona facilities, announced in 2020 in anticipation of federal support and expanded substantially after the CHIPS and Science Act awards. The sample window, however, understates the losses from a U.S. location. Fab 14 continues to operate beyond 2015, its technology continues to mature, and the annual profit losses from a U.S. location are largest in exactly these later years. Because the subsidies are received up front while the losses accumulate for as long as the facility operates, continued U.S. operations without additional later-stage incentives could be a concern: TSMC could face incentives to wind down U.S. operations and to direct utilization and upgrades toward its capacity in Taiwan instead. Policymakers' perceived benefit of current policies may be overly optimistic for this reason, and similar considerations apply to TSMC's Arizona fabs if these location-based cost differences persist.

The composition of the subsidy package sharpens this concern. The ITC accounts for $2.0 billion of the $2.7 billion in total incentives, and it is by statute proportional to up-front investment: it pays out as qualified property is placed in service, regardless of how long or how intensively the facility subsequently operates. The largest instrument in the package is therefore also the one most tightly tied to the participation margin and least able to reach the continuation margin. The 2025 increase in the credit rate from 25% to 35% raises this share further, expanding the package in precisely the dimension that our results suggest is not the binding one.[16]

D Adding U.S. Import Tariffs to the Current Policy Regime

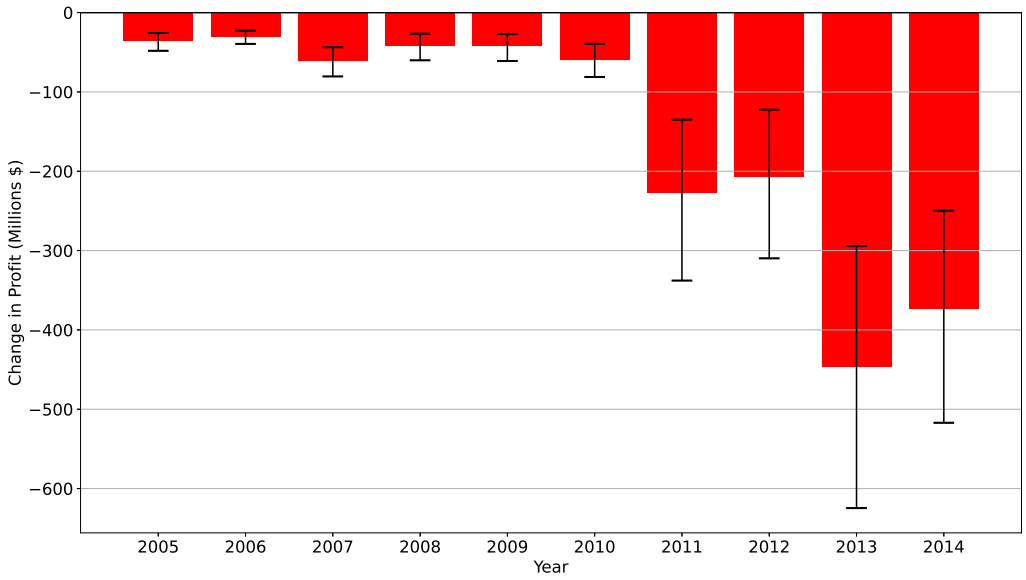
One additional policy option that has recently attracted significant attention is the use of import tariffs. The U.S. government has discussed tariff rates of 25% or higher (York and Durante, 2025), while actual tariff rates have ranged from 10% to 14% over 2025–2026 (McClelland and Wong, 2026). In this simulation, we examine the potential impact of import tariffs in addition to the current policy regime. Import tariffs function as a percentage levy on U.S. buyers purchasing imported goods, effectively reducing their willingness to pay when the manufacturer is located outside the U.S. Since this reduction applies to the total price—rather than just the

[16] Appendix C.3 shows that placing Fab 14 on a more aggressive upgrade schedule could require an additional $3.4 billion of investment in the U.S. and would draw roughly $1.2 billion of additional incentives, but would raise the firm's simulated profit within our window by only $0.5 billion, so even an instrument that does reward later investment leaves about half of the incremental spending uncovered. Addressing the continuation margin may therefore require policies that more directly offset the higher marginal cost of producing in the U.S.

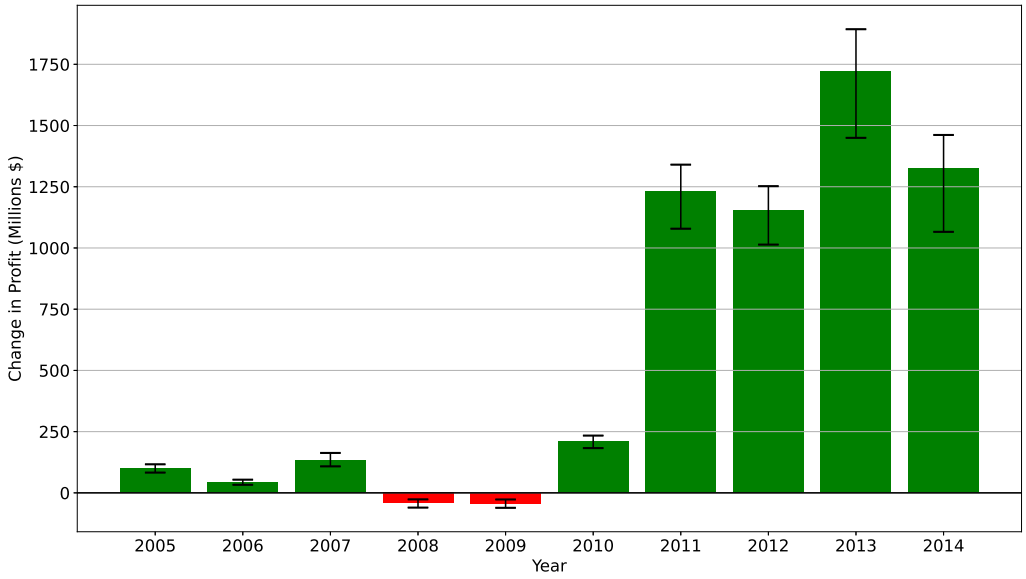
profit margin—its effect is likely to be stronger than that of an equivalent corporate tax cut. However, due to uncertainty around the precise magnitude of the effect, we conservatively model tariffs by reducing U.S. buyers' willingness to pay (v_{igt}) by 10% of the common marginal cost for fabs located outside the U.S.[17]

Fig. 3: Annual Change in TSMC Profit: Fab 14 Relocated to the U.S. vs. Remaining in Taiwan

(A) Current Regime: U.S. Export Controls + CN Tax Exemptions



(B) Adding U.S. Import Tariffs to Current Regime



Notes: Each figure plots the change in profit of relocating Fab 14 to the U.S. relative to letting Fab 14 remain in Taiwan under the respective policy regime. Red bars indicate a decrease in profit, while green bars indicate an increase. The 90% confidence intervals are shown using error bars—horizontal caps above and below a vertical line—on each bar.

[17] We assume that under this adjustment, buyer valuations are still high enough such that all firms still submit bids.

The results are revealing in several respects. First, tariffs have a limited effect when Fab 14 is located in Taiwan. Compared to the current policy regime, adding tariffs results in no visible change in TSMC's profit under our baseline 10% calibration. Even increasing the calibration to 20% or 25% leads to only a modest decline in profit of about $2 billion. This result is quite intuitive within our model. During our sample period, only 6% of contract manufacturing orders were fulfilled by fabs located in North America due to minimal domestic capacity. As a result, buyers had limited alternatives within the U.S. and, even if they preferred to avoid tariffs, had little choice but to import. Our model assumes inelastic demand on the buyer side, which is a reasonable assumption for much of the market between 2004 and 2015.[18] According to IC Insights (2022), semiconductor content accounted for only 21 – 26% of the total value of electronic systems during that period. This suggests that increases in semiconductor costs from tariffs would represent a relatively small portion of the total cost of electronics, which is unlikely to induce a significant reduction in downstream demand.[19] As a result, buyers bear the full cost of tariffs, and tariffs have minimal impact on manufacturers located outside the U.S., conditional on the continued lack of U.S. capacity.

When Fab 14 is relocated to the U.S. under a tariff regime, profit increases substantially relative to the Taiwan location. Specifically, TSMC earns $70.9 billion in profit under the combined tariffs, export controls, and Chinese tax exemptions scenario when Fab 14 is relocated to the U.S., $5.9 billion more than the $65.0 billion earned if Fab 14 remains in Taiwan. What accounts for this gain? Figure 3B plots the annual profit difference between Fab 14 in Taiwan and in the U.S. under the tariff scenario, revealing that the gains occur almost entirely in the later years, when Fab 14's technology matures. At that stage, the U.S.-based fab captures a significant share of domestic demand for mature nodes—demand that it previously had to compete for globally. In this context, tariffs counterbalance the competitive disadvantages for U.S. demand typically faced by fabs producing mature technologies in high-cost locations like the U.S., leading to the opposite patterns observed in Figure 3B compared to Figure 3A. This tariff-induced reversal of location incentives is not unique to our setting: Flaaen et al. (2020)

[18] Appendix C.1C provides further evidence that demand is inelastic during this period using an instrumental-variables approach. The resulting estimates are statistically indistinguishable from perfectly inelastic demand and have tight confidence intervals.

[19] However, downstream demand could be more elastic for high-end products such as flagship smartphones or standalone GPUs used in gaming or High-Performance Computing (HPC) clusters, where the semiconductor share of system value is higher.

document that U.S. tariffs on washing machines induced foreign manufacturers to relocate production to the U.S., with the accompanying price increases borne by domestic consumers, a pattern that parallels both our relocation and our welfare results.

Welfare Analysis While our analysis above focuses on TSMC's opportunity cost of operating Fab 14 in the U.S. rather than in Taiwan, our model and estimates are also informative about the welfare consequences of counterfactual policies and how these consequences interact with TSMC's location decision. From a policymaker's perspective, tariffs are particularly interesting because they affect the payoffs of both U.S. buyers and manufacturers. Table 8 summarizes and decomposes the welfare effects of the tariff policy on U.S. chip buyers.

When Fab 14 is located in Taiwan, adding the import tariffs specified in Section 7D to the current policy regime reduces U.S. chip buyer surplus by $48.9 billion. Of this amount, $44.0 billion reflects tariff payments, while $4.9 billion reflects losses from buyers reallocating orders toward less cost-efficient U.S. manufacturers once tariffs raise the effective cost of purchasing from overseas manufacturers. The loss from the latter channel is relatively limited because U.S. manufacturing capacity is limited, although the policy generates a substantial transfer of surplus from U.S. buyers to the U.S. government.

When Fab 14 is relocated to the U.S., adding the same import tariffs reduces U.S. chip buyer surplus by $47.7 billion, approximately $1.0 billion less than when Fab 14 remains in Taiwan. The primary reason is that the additional U.S. capacity allows more U.S. buyers to source domestically and thereby avoid tariffs. As a result, tariff payments fall to $36.0 billion, $8.0 billion less than in the Taiwan scenario. At the same time, however, the additional U.S. capacity is less cost-efficient than overseas capacity. Tariffs therefore induce greater reallocation of orders toward less cost-efficient U.S. production, increasing the associated buyer-surplus loss by $7.0 billion relative to the Taiwan scenario. Because tariff payments accrue to the U.S. government and are therefore transfers within the U.S. economy rather than pure welfare losses, a measure of net U.S. welfare loss should focus on the efficiency loss from reallocating production toward less cost-efficient manufacturers. By this measure, relocating Fab 14 to the U.S. increases the net welfare loss by $7.0 billion, which exceeds the $5.9 billion increase in TSMC's profit generated by relocation.

In summary, our analysis indicates that under the tariff scenario, relocating Fab 14 to the

U.S. becomes profitable for TSMC relative to keeping it in Taiwan. Even at the modest tariff rate simulated here, the additional U.S. contracts and $5.9 billion increase in profit from a domestic presence more than offset the additional $1.2 billion investment cost. However, the same policy induces substantial losses for U.S. chip buyers relative to a no-tariff scenario, and the associated efficiency loss from reallocating production toward less cost-efficient domestic capacity exceeds the gains captured by the relocating manufacturer.

Table 8: Changes in U.S. Chip Buyer Surplus from Import Tariffs

	U.S. Import Tariffs + Current Policies vs. Current Policies		
	Δ U.S. Buyer Surplus	Due to U.S. Tariff Payments	Due to Order Reallocation to Less Cost-Efficient Manufacturers
	(1)	(2)	(3)
Fab 14 in Taiwan	−$48.9B	−$44.0B	−$4.9B
	[−53.1, −44.7]	[−47.7, −40.3]	[−5.8, −4.1]
Fab 14 relocated to the U.S.	−$47.7B	−$36.0B	−$11.8B
	[−51.8, −43.6]	[−39.3, −32.6]	[−13.0, −10.6]
Difference	+$1.0B	+$8.0B	−$7.0B
	[0.9, 1.2]	[7.6, 8.5]	[−7.5, −6.6]

Notes: The table reports changes in surplus for U.S. chip buyers from adding the U.S. import tariffs specified in Table 7 to the current policy regime, which includes U.S. export controls and Chinese tax exemptions. Columns (2) and (3) decompose the total change in buyer surplus into tariff payments and the additional welfare losses arising from contracting with less cost-efficient manufacturers. The first two rows report the results when TSMC Fab 14 is located in Taiwan and when it is relocated to the U.S., respectively. The final row reports the difference between the relocation and Taiwan scenarios. All remaining notes from Table 7 apply.

8 Limitations

While our model and estimation results reasonably capture the variations in both variable and fixed manufacturing costs with plausible estimates and signs, our approach is not without limitations. We highlight several key limitations here. These limitations are unlikely to undermine the counterfactual comparisons of manufacturer payoffs across locations and policy regimes reported above, which are the primary focus of this paper. They bear more directly on the welfare analysis of Table 8, and on broader questions we leave to future work, such as the full societal welfare effects of industrial policies.

First, our model does not account for potential long-term relationships between buyers and manufacturers, where buyers might need to commit to contracts many quarters in advance to secure capacity, or even co-invest in the R&D or capacity of manufacturers. If such relationships were extensive, the contracting process would need to be fully dynamic, rather than the static process assumed in our model, which treats current capacity across manufacturers in each period as given. However, it is reasonable to believe that long-term relationships are not widespread.

For example, an interview with TSMC's founder Morris Chang (Acquired, 2025) noted that buyers often backed out of contracts last-minute and did not commit to capacity investments, making capacity investment a risky business for contract manufacturers. As Morris Chang stated,"It's our money, and it's only their words."

Second, we have assumed that marginal cost decreases over time exogenously. In reality, however, an important driver of the decline in marginal cost is the volume of chips a manufacturer has produced, meaning that marginal cost should ideally be a function of the number of previous orders. This also implies that pricing strategies must consider how prices affect order quantity and future marginal cost.[20] As a result, the equilibrium price would need to be fully dynamic, accounting for these factors. We do not pursue this approach because it would make the modeling intractable. As a practical matter, observed prices in the data do not allow one to distinguish between a pricing schedule under exogenously decreasing marginal cost versus a pricing schedule resulting from incorporating the effect of volume on marginal cost and endogenously changing marginal cost due to volume.

Furthermore, our assumption that buyer valuations are such that the manufacturer's auction participation constraints never bind also has implications for how we interpret our estimated marginal cost parameters. In particular, this assumption may generate an upward bias in our firm-specific costs for firms with higher marginal costs. Under this assumption, we rationalize firms losing auctions only when they have higher marginal costs than competitors, rather than choosing not to bid because their costs exceed the buyer's valuation. The latter would allow for firms with low marginal costs in absolute terms to not participate because buyers' valuations are also low in absolute terms. We do not believe this assumption greatly affects our counterfactual assumptions in this paper, as we focus on profits of the lowest marginal cost firms (i.e., TSMC). This assumption also means that we cannot identify the level of buyer valuations. The welfare analysis of Table 8 therefore reports changes in buyer surplus rather than levels: because order quantities are inelastic in our main specification, the change in a buyer's surplus is the change in the price it pays times the quantity it orders, which does not require knowing the level of its valuation. Statements about the level of buyer welfare, or about the total surplus the market generates, remain outside what our estimates identify.

[20]The pricing strategy that incorporates this thinking is called learning curve pricing, which involves setting lower-than-profit-maximizing prices initially for a new technology and committing to a pre-determined declining price schedule for the technology.

Lastly, due to the lack of granular data within each fab, we assume that each fab manufactures only one technology generation at a time.

9 Conclusion

Our analysis shows that industrial policies promoting domestic self-reliance operate through two channels with different time profiles. Subsidies at the scale of the CHIPS and Science Act are sufficient to cover the up-front investment cost of a U.S. location and can rationalize a manufacturer's decision to build. They do not address the flow-profit penalty of operating in the U.S., which grows as a facility's technology matures and which no front-loaded, lump sum payment reaches. Policymaker beliefs about the later-stage returns to subsidizing up-front investment costs may therefore be overly optimistic. This concern applies with particular force to the investment tax credit, the largest component of the current incentive package, which is by design proportional to up-front investment and was expanded further in 2025. Promoting durable domestic production may require commitment to later-stage instruments, such as subsidies tied to continued operations or upgrades. Among the instruments we study, only tariffs operate on the per-period margin that matters for retention, but our welfare calculations show that U.S. chip buyer losses from tariffs exceed the relocated manufacturer's gains.

We view this paper as a first step toward designing and evaluating industrial policy in the semiconductor industry, and we limit the scope of our exercise. We hold the observed configuration of firms, capacities, and technologies fixed, so our counterfactuals do not capture entry, exit, or investment responses by TSMC's competitors, nor TSMC's own re-optimization of its upgrade path. Our welfare analysis covers buyers and sellers in the contract manufacturing market but not downstream consumers or broader objectives, such as supply chain resilience, that motivate these policies. Extending the analysis to endogenous investment responses and to these broader objectives is a natural direction for future work.

References

Acquired. TSMC founder Morris Chang. Podcast, 2025. URL `https://www.acquired.fm/episodes/tsmc-founder-morris-chang`. Accessed: February 20, 2025.

Jason Allen, Robert Clark, and Jean-François Houde. Search frictions and market power in negotiated-price markets. *Journal of Political Economy*, 127(4):1550–1598, August 2019.

Richard Baldwin and Paul R. Krugman. Market access and international competition: A simulation study of 16k random access memories. In Robert C. Feenstra, editor, *Empirical Methods for International Trade*, pages 171–197. MIT Press, Cambridge, MA, 1988.

Panle Jia Barwick, Myrto Kalouptsidi, and Nahim Bin Zahur. Industrial policy implementation: Empirical evidence from China's shipbuilding industry. *Review of Economic Studies*, 92(6): 3611–3648, 2025.

Jon Bateman. US-China technological "decoupling": A strategy and policy framework. 2022. URL `https://carnegieendowment.org/files/Bateman_US-China_Decoupling_final.pdf`. Accessed on March 19, 2024.

Chad P Bown and Dan Wang. Semiconductors and modern industrial policy. *Journal of Economic Perspectives*, 38(4):81–110, 2024.

David M Byrne, Brian K Kovak, and Ryan Michaels. Quality-adjusted price measurement: A new approach with evidence from semiconductors. *Review of Economics and Statistics*, 99 (2):330–342, 2017.

Jaedo Choi and Andrei A Levchenko. The long-term effects of industrial policy. *Journal of Monetary Economics*, 152:103779, 2025.

CNBC. Chipmakers get bigger tax credits in Trump's latest big beautiful bill. CNBC, jul 2025. URL `https://www.cnbc.com/2025/07/02/chipmakers-get-bigger-tax-credits-in-trumps-latest-big-beautiful-bill.html`. Accessed July 21, 2025.

Chiara Criscuolo, Ralf Martin, Henry G Overman, and John Van Reenen. Some causal effects of an industrial policy. *American Economic Review*, 109(1):48–85, 2019.

Design & Reuse. TSMC continues to dominate the worldwide foundry market. 2018. URL `https://www.design-reuse.com/news/43943/2017-foundries-marketshares.html`. Accessed on January 28, 2024.

Bilge Erten, Joseph E Stiglitz, and Eric Verhoogen. Employment impacts of the CHIPS Act. *Brookings Papers on Economic Activity*, 2025(1):1–63, 2025.

Pablo D. Fajgelbaum, Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal. The return to protectionism. *The Quarterly Journal of Economics*, 135(1):1–55, 2020.

Aaron Flaaen, Ali Hortaçsu, and Felix Tintelnot. The production relocation and price effects of US trade policy: The case of washing machines. *American Economic Review*, 110(7): 2103–2127, 2020.

Aaron L Friedberg and Charles W Boustany Jr. Partial disengagement: A new US strategy for economic competition with China. *The Washington Quarterly*, 43(1):23–40, 2020.

Ronald L Goettler and Brett R Gordon. Does AMD spur Intel to innovate more? *Journal of Political Economy*, 119(6):1141–1200, 2011.

Pinelopi K Goldberg, Réka Juhász, Nathan J Lane, Giulia Lo Forte, and Jeff Thurk. Industrial policy in the global semiconductor sector. 2024. NBER Working Paper 32651.

Pengfei Han, Wei Jiang, and Danqing Mei. Mapping U.S. – China technology decoupling: Policies, innovation, and firm performance. *Management Science*, page mnsc.2022.02057, February 2024.

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.

Yuka Hayashi and Asa Fitch. U.S. tightens curbs on AI chip exports to China, widening rift

with Nvidia and Intel. *Wall Street Journal*, October 2023.

Joel Hruska. How are process nodes defined? *ExtremeTech*, Feb 2021. URL https://www.extremetech.com/computing/296154-how-are-process-nodes-defined. Accessed on January 4, 2024.

IC Insights. Value of semiconductor content in electronic systems worldwide from 1999 to 2021. Statista, January 2022. URL https://www.statista.com/statistics/1287839/value-semiconductor-content/. Accessed July 24, 2025.

Mitsuru Igami. Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998. *Journal of Political Economy*, 125(3):798–847, 2017.

Mitsuru Igami and Kosuke Uetake. Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2016. *The Review of Economic Studies*, 87(6): 2672–2702, 2020.

Industry and Security Bureau. Addition of entities to the entity list, revision of entry on the entity list, and removal of entities from the entity list. *Federal Register*, December 2020.

Industry and Security Bureau. Implementation of additional export controls: Certain advanced computing and semiconductor manufacturing items; supercomputer and semiconductor end use; entity list modification. *Federal Register*, October 2022.

Industry and Security Bureau. Implementation of additional export controls: Certain advanced computing items; supercomputer and semiconductor end use; updates and corrections. *Federal Register*, October 2023.

Brian W. Jacobs, Vinod R. Singhal, and Xinrui Zhan. Stock market reaction to global supply chain disruptions from the 2018 US government ban on ZTE. *Journal of Operations Management*, 68(8):903–927, December 2022.

Réka Juhász. Temporary protection and technological development: Evidence from the napoleonic blockade. *American Economic Review*, 108(11):3339–3376, 2018.

Réka Juhász, Nathan Lane, and Dani Rodrik. The new economics of industrial policy. *Annual Review of Economics*, 16:213–242, 2024.

Myrto Kalouptsidi. Time to build and fluctuations in bulk shipping. *American Economic Review*, 104(2):564–608, 2014.

Myrto Kalouptsidi. Detection and consequences of industrial subsidies: The case of shipbuilding. *The Review of Economic Studies*, 85(2):1111–1158, 2018.

Saif M. Khan. U.S. semiconductor exports to China: Current policies and trends. October 2020. URL https://cset.georgetown.edu/publication/u-s-semiconductor-exports-to-china-current-policies-and-trends/. Accessed on March 18, 2024.

Dan Kim and John VerWey. The potential impacts of the Made in China 2025 Roadmap on the integrated circuit industries in the U.S., EU and Japan. 2019. URL https://www.ssrn.com/abstract=3433844. Working paper.

Yunmi Kong. Identification of english auctions when losing entrants are not observed. *International Journal of Industrial Organization*, 85:102875, 2022. ISSN 0167-7187. doi: https://doi.org/10.1016/j.ijindorg.2022.102875. URL https://www.sciencedirect.com/

science/article/pii/S0167718722000510.

Nathan Lane. Manufacturing revolutions: Industrial policy and industrialization in South Korea. *The Quarterly Journal of Economics*, 140(1):279–349, 2025.

Robert McClelland and John Wong. TPC Tariff Tracker. Tax Policy Center, July 2026. URL https://taxpolicycenter.org/features/tracking-trump-tariffs. Accessed August 9, 2026.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026, 1989.

Weiting Miao. Technology rivalry and resilience under trade disruptions: The case of semiconductor foundries. 2024. Working Paper.

Nathan H. Miller, Matthew Osborne, Gloria Sheu, and Gretchen Sileo. Technology and market power in the cement industry. Working paper, December 2025. URL https://www.nathanhmiller.org/cementmarkups.pdf.

Ministry of Finance of the People's Republic of China, State Taxation Administration of the People's Republic of China, National Development and Reform Commission (NDRC), and Ministry of Industry and Information Technology of the People's Republic of China. 关于促进集成电路产业和软件产业高质量发展企业所得税政策的公告, 2020. URL https://www.gov.cn/zhengce/zhengceku/2020-12/17/content_5570401.htm. Accessed on December 31, 2023.

National Information Standards Organization. ANSI/NISO Z39.104-2022, CRediT, Contributor Roles Taxonomy. Standard ANSI/NISO Z39.104-2022, National Information Standards Organization (NISO), Baltimore, MD, 2022.

OECD. Measuring distortions in international markets: The semiconductor value chain. 2019. URL https://www.oecd-ilibrary.org/content/paper/8fe4491d-en. Accessed on December 31, 2023.

Ariel Pakes and David Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057, 1989.

Reuters. China sets up third fund with $47.5 bln to boost semiconductor sector. *Reuters*, May 2024. URL https://www.reuters.com/technology/china-sets-up-third-fund-with-475-bln-boost-semiconductor-sector-2024-05-27/. Accessed July 21, 2025.

Stephen P. Ryan. The costs of environmental regulation in a concentrated industry. *Econometrica*, 80(3):1019–1061, 2012.

Semiconductor Engineering. Nodes. n.d. URL https://semiengineering.com/knowledge_centers/manufacturing/process/nodes/. Accessed on January 4, 2024.

Semiconductor Industry Association. Beyond borders: the global semiconductor value chain. 2016. URL https://www.semiconductors.org/wp-content/uploads/2018/06/SIA-Beyond-Borders-Report-FINAL-June-7.pdf. Accessed on January 3, 2024.

Semiconductor Industry Association. U.S. needs greater semiconductor manufacturing incentives, 2020. URL https://www.semiconductors.org/wp-content/uploads/2020/07/U.S.-Needs-Greater-Semiconductor-Manufacturing-Incentives-Infographic1.pdf. Ac-

cessed on December 30, 2023.

Semiconductor Industry Association. SIA whitepaper: Taking stock of China's semiconductor industry. 2021. URL https://www.semiconductors.org/wp-content/uploads/2021/07/Taking-Stock-of-China%E2%80%99s-Semiconductor-Industry_final.pdf. Accessed on December 31, 2023.

Semiconductor Industry Association. SIA applauds House passage of CHIPS Act, urges President to sign bill into law, 2022. URL https://www.semiconductors.org/sia-applauds-house-passage-of-chips-act-urges-president-to-sign-bill-into-law/. Accessed on December 30, 2023.

Ana Swanson. U.S. issues final rules to keep chip funds out of China. *New York Times*, September 2023. URL https://www.nytimes.com/2023/09/22/us/politics/us-final-rules-chip-makers-china.html. Accessed on December 30, 2023.

Taiwan Semiconductor Manufacturing Company. TSMC announces intention to build and operate an advanced semiconductor fab in the United States, 2020. URL https://pr.tsmc.com/system/files/newspdf/THGOANPGTH/NEWS_FILE_EN.pdf. Accessed on March 20, 2024.

The State Council of the People's Republic of China. 国务院关于印发《中国制造 2025》的通知. 2015. URL https://www.gov.cn/zhengce/content/2015-05/19/content_9784.htm. Accessed on January 1, 2024.

Jeff Thurk. Outsourcing, firm innovation, and industry dynamics in the production of semiconductors. 2022. Working Paper.

U.S. Chamber of Commerce. Understanding U.S.-China decoupling: Macro trends and industry impacts. February 2021. URL https://www.uschamber.com/assets/archived/images/024001_us_china_decoupling_report_fin.pdf. Accessed on January 1, 2024.

U.S. Government Accountability Office. Semiconductor supply chain: Policy considerations from selected experts for reducing risks and mitigating shortages, 2022. URL https://www.gao.gov/assets/gao-22-105923.pdf. Accessed on March 19, 2024.

U.S. Senate Committee on Commerce, Science, and Transportation. CHIPS and Science Act of 2022 Division A summary - CHIPS and ORAN investment, 2022a. URL https://www.commerce.senate.gov/services/files/2699CE4B-51A5-4082-9CED-4B6CD912BBC8. Accessed on December 30, 2023.

U.S. Senate Committee on Commerce, Science, and Transportation. CHIPS and Science Act of 2022 section-by-section summary, 2022b. URL https://www.commerce.senate.gov/services/files/1201E1CA-73CB-44BB-ADEB-E69634DA9BB9. Accessed on December 30, 2023.

Antonio Varas and Raj Varadarajan. How restrictions to trade with China could end US leadership in semiconductors. March 2020. URL https://web-assets.bcg.com/img-src/BCG-How-Restricting-Trade-with-China-Could-End-US-Semiconductor-Mar-2020_tcm9-240526.pdf. Accessed on January 1, 2024.

Antonio Varas, Raj Varadarajan, Jimmy Goodrich, and Falan Yinug. Government incentives and US competitiveness in semiconductor manufacturing. September 2020. URL https://www.semiconductors.org/wp-content/uploads/2020/09/Government-Incentives-and-US-Competitiveness-in-Semiconductor-Manufacturing-Sep-2020.pdf. Accessed on March 18, 2024.

John VerWey. Chinese semiconductor industrial policy: Past and present. *United States International Trade Commission Journal of International Commerce and Economics*, 2019.

Kaiqi Wang. 大基金二期投资图鉴，虎年大步快跑起来. 全球半导体观察, March 2022. URL https://www.dramx.com/News/IC/20220307-30936.html. Accessed on December 31, 2023.

Xinhua News Agency. 国务院印发《国家集成电路产业发展推进纲要》. 2014. URL https://www.gov.cn/xinwen/2014-06/24/content_2707281.htm. Accessed on January 1, 2024.

Erica York and Alex Durante. Tracking the economic impact of the Trump tariffs. *Tax Foundation*, jul 2025. URL https://taxfoundation.org/research/all/federal/trump-tariffs-trade-war/. Last updated July 17, 2025; Accessed July 21, 2025.

Yuanchuang Zhao. 大基金一期的投资成果回顾. 半导体行业观察, March 2021. URL http://www.semiinsights.com/s/electronic_components/23/41742.shtml. Accessed on December 31, 2023.

Appendices

A Additional Industry and Policy Background

A.1 Policy Details

This appendix provides additional institutional details on the semiconductor policies summarized in the main text, with particular attention to their implementation, eligibility restrictions, and broader international context.

United States. The CHIPS and Science Act of 2022 encompasses a broader set of programs than the manufacturing incentives emphasized in our counterfactuals. In total, the legislation authorized $52.7 billion for semiconductor-related initiatives, including manufacturing incentives, R&D, workforce development, and public–private partnerships (U.S. Senate Committee on Commerce, Science, and Transportation, 2022a). An important feature of the manufacturing incentives is the accompanying set of geographic guardrails. Recipients of CHIPS financial assistance face restrictions on materially expanding advanced semiconductor manufacturing capacity for ten years in countries designated as presenting national security concerns, including China, Iran, Russia, and North Korea (Swanson, 2023). These provisions therefore tie domestic manufacturing support to restrictions on firms' subsequent capacity investments abroad.

U.S. export controls impose a separate set of restrictions on semiconductor-related transactions with China (Khan, 2020; Industry and Security Bureau, 2020, 2022, 2023). The controls operate through several mechanisms. List-based restrictions cover specified advanced chips and semiconductor manufacturing equipment and require exporters to obtain licenses for covered transactions. Entity- and end-use-based restrictions impose additional requirements on transactions involving designated Chinese firms or specified applications. Major Chinese semiconductor manufacturers, including SMIC, have been subject to such restrictions. Spanning a broader set of technologies, these controls generally assume a default position of denying export licenses. Unlike manufacturing subsidies, which affect the returns to locating capacity in the U.S., these controls directly constrain firms' access to particular customers and markets.

The policy emphasis on domestic capacity reflects longstanding concerns about the decline in the U.S. share of global semiconductor manufacturing. In motivating the CHIPS legislation, the U.S. Senate Committee on Commerce, Science, and Transportation highlighted the decline in the domestic share of semiconductor production from approximately 37% in the 1990s to approximately 12%, together with increasing government support for semiconductor manufacturing abroad. Thus, although the legislation also supports research and workforce development, its manufacturing provisions explicitly seek to alter the geographic distribution of production capacity.

China. China's recent semiconductor policies build on a longer-running effort to develop a domestic semiconductor industry (VerWey, 2019). These efforts intensified in 2014 with the

release of the Guidelines to Promote the National Integrated Circuit Industry (IC Guidelines), which established ambitious objectives for expanding China's semiconductor capabilities (Xinhua News Agency, 2014; Semiconductor Industry Association, 2021). The Made in China 2025 plan, released the following year, further emphasized semiconductor self-sufficiency and articulated an aspirational target of producing 70% of domestic semiconductor demand within China by 2025 (The State Council of the People's Republic of China, 2015; Semiconductor Industry Association, 2021).

A central institutional mechanism for pursuing these objectives has been the National Integrated Circuit Industry Development Investment Fund, commonly referred to as the "Big Fund." Rather than operating primarily as a conventional grant program, the Big Fund provides state-backed equity financing to semiconductor firms. By mid-2021, approximately 70% of its cumulative investment had been directed toward semiconductor manufacturing (Semiconductor Industry Association, 2021). Funding has primarily supported Chinese-owned firms expanding manufacturing capacity within China (Zhao, 2021; Wang, 2022). This structure differs from broad-based R&D support because financing is directly tied to firms and investments intended to expand domestic production capabilities.

China also uses several complementary instruments. Corporate tax incentives vary with manufacturing technology: firms operating at more advanced technology nodes qualify for longer exemption periods (Ministry of Finance of the People's Republic of China et al., 2020). In addition, semiconductor manufacturers may receive financing from state-owned financial institutions at below-market interest rates. The magnitude and allocation of such credit support are considerably less transparent than formal fiscal programs, making them difficult to quantify systematically (OECD, 2019). These policies collectively combine equity financing, tax preferences, and potentially preferential credit to support domestic capacity expansion.

Other countries and regions. Manufacturing incentives are not unique to the U.S. and China. South Korea, Singapore, Japan, Taiwan, and European countries have all used combinations of grants, tax incentives, concessional financing, infrastructure support, and other measures to encourage semiconductor investment (Semiconductor Industry Association, 2020). Historically, however, the scale of direct manufacturing support in these economies has generally been smaller than the major programs introduced in China and, more recently, the U.S. For

example, estimates for 2000–2020 place Taiwan's manufacturing grants and subsidies below $1 billion and South Korea's at approximately $7–10 billion (Semiconductor Industry Association, 2020). These programs provide useful international context, but our counterfactual analysis focuses on the policy instruments most central to the recent U.S.–China policy environment.

A.2 Calibration of Policy Counterfactuals

This appendix records how the policies simulated in Tables 6 and 7 map into model objects.

U.S. capacity subsidies Column (5) of Table 6 sets the direct subsidy at 10% of total investment, matching TSMC Arizona's $6.6 billion award on a $65 billion investment. Column (6) applies the ITC at 25% of qualified investment, the rate established under Section 48D of the Internal Revenue Code in 2022 and the rate under which TSMC Arizona's award was made; the rate was increased to 35% for property placed in service after December 31, 2025. The credit offsets tax liability dollar for dollar and is effectively refundable through the statute's elective payment provision, so its value does not depend on the firm's U.S. tax liability. Both instruments are treated as lump sums and reported at face value. The face values abstract from two tax offsets: the credit reduces the depreciable basis of the property, forfeiting future depreciation deductions worth up to the credit multiplied by the combined federal and state corporate tax rate of 26% (a 21% federal rate and a 5% state rate; Arizona's rate is 4.9%), and the direct subsidy is generally taxable as corporate income.

U.S. export controls The U.S. export control policy issued in 2022 restricted exports of chips produced with technology nodes at 16/14 nanometers or below (Industry and Security Bureau, 2022). The rule took effect when TSMC's most advanced technology in volume production was 3nm, which reflects a four-generation gap, with 10/12nm and 5/7nm in between. Our simulations therefore prohibit TSMC from contracting with mainland Chinese buyers for technology generations within four generations of its most advanced technology in each period, anchoring the restriction to the moving frontier rather than to a fixed node.

Chinese tax exemptions The exemptions described in Ministry of Finance of the People's Republic of China et al. (2020) reduce the standard 25% corporate income tax rate for chipmakers located in mainland China to between 0% and 12.5%, depending on a firm's technology

portfolio and the number of years since the policy announcement. In our model, profit margins are determined by the gap between the lowest and second-lowest marginal costs, so these exemptions act as marginal cost subsidies for fabs operating in mainland China. Because it is unclear how much of the tax savings are passed through to buyers in the form of lower bid prices, we apply a conservative calibration and reduce the region-specific marginal cost in China (c_{CN}) by 5% of the technology-generation- and time-specific common marginal cost (c_{gt}). All other cost and demand parameters are held at their estimated values.

U.S. import tariffs A tariff rate of 25% or higher has been discussed by the U.S. government (York and Durante, 2025) and actual rates range between 10–14% over 2025–2026 (McClelland and Wong, 2026). Tariffs act as a percentage levy on U.S. buyers purchasing imported goods, reducing their willingness to pay when the manufacturer is located outside the U.S. Because this reduction applies to the total price rather than only to the profit margin, its effect should be stronger than that of an equivalent corporate tax reduction. The exact magnitude is uncertain, so we conservatively reduce U.S. buyers' willingness to pay (v_{igt}) by 10% of the common marginal cost for fabs located outside the U.S.

B Model Details

B.1 Equilibrium Expressions

With i.i.d. extreme-value cost shocks, the equilibrium of the bidding game in Section 4 admits closed-form expressions. Each manufacturer j wins buyer i's contract with probability

$$P_{j,igt}[j=(1)|\boldsymbol{s}_{gt})] = \frac{exp(-(c_{gt}^{j} + c_{j,igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{gt}^{j}\})(\lambda + \kappa_{igt}))/(c_{gt}\sigma_{\omega}))}{\sum_{j' \in \mathcal{J}_{igt}} exp(-(c_{gt}^{j'} + c_{j',igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{gt}^{j'}\})(\lambda + \kappa_{igt}))/(c_{gt}\sigma_{\omega}))}. \tag{B.1}$$

v does not appear in the above expression because it is common across all participating manufacturers with available capacity. If manufacturer j is the winner, then its expected price

is

$$E[p_{igt}^{(1)}(\boldsymbol{s}_{gt})] = \sum_{k \in \mathcal{J}_{igt} \setminus j} \Bigg(\frac{exp(-(c_{gt}^{k} + c_{k,igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{gt}^{k}\})(\lambda + \kappa_{igt}))/(c_{gt}\sigma_{\omega}))}{\sum_{j' \in \mathcal{J}_{igt} \setminus j} exp(-(c_{gt}^{j'} + c_{j',igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{gt}^{j'}\})(\lambda + \kappa_{igt}))/(c_{gt}\sigma_{\omega}))} \tag{B.2}$$

$$\times E[(c_{gt}^{(2)} + c_{(2),igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{gt}^{(2)}\})(\lambda + \kappa_{igt}) + \omega_{(2),igt} + \mathbf{1}(geog_{igt} \in \{geog_{gt}^{(2)}\})\kappa_{igt})] \Bigg)$$

B.2 Additional Parameterization Details

Buyer location shares We parameterize the share of contracts from buyers in these regions, $d_{gt}(geog_i|\cdot)$,t as follows:

$$d_{gt}(geog_i = CN|g_{it}, t) = \frac{exp(l_0^{CN} + l_1^{CN}(g_t^{max} - g_{it}) + l_2^{CN}t)}{1 + \sum_{geog}^{\{CN,NA\}} exp(l_0^{geog} + l_1^{geog}(g_t^{max} - g_{it}) + l_2^{geog}t)} \tag{B.3}$$

$$d_{gt}(geog_i = NA|g_{it}, t) = \frac{exp(l_0^{NA} + l_1^{NA}(g_t^{max} - g_{it}) + l_2^{NA}t)}{1 + \sum_{geog}^{\{CN,NA\}} exp(l_0^{geog} + l_1^{geog}(g_t^{max} - g_{it}) + l_2^{geog}t)} \tag{B.4}$$

$$d_{gt}(geog_i = ROW|g_{it}, t) = \frac{1}{1 + \sum_{geog}^{\{CN,NA\}} exp(l_0^{geog} + l_1^{geog}(g_t^{max} - g_{it}) + l_2^{geog}t)} \tag{B.5}$$

These functional forms are informed by industry knowledge: North American buyers on average demand more cutting edge technologies, and the share of Chinese buyers has steadily increased while the share of North American buyers has steadily decreased over time since the 2000s. The parameter l_0^{CN} is identified by the average revenue share from Chinese buyers across manufacturers. The parameter l_1^{CN} is identified by comparing the revenue share from Chinese buyers across manufacturers located in the same region but with different technology portfolios (e.g., TSMC vs. UMC, SMIC vs. Hua Hong). The parameter l_2^{CN} is identified by tracking changes in the share of revenue from Chinese buyers for the same manufacturers over time (e.g., time series from TSMC, UMC, SMIC). The identification of l_0^{NA}, l_1^{NA}, and l_2^{NA} follows the same logic.[21]

IDM capacity available for contract manufacturing For an IDM, the total capacity available for contract manufacturing at facility k in period t is $m_{jt}^k Q_{jt}^k$, where m_{jt}^k equals $min\{max\{0, \tilde{m}_{jt}^k\}, 1\}$ and $\tilde{m}_{jt}^k$ is drawn i.i.d. from a normal distribution $N(0, \sigma_{\tilde{m}})$ with $\sigma_{\tilde{m}} = 0.05$.

[21] $t = 2009Q4$ is normalized to a value of zero in the above equations.

C Further Extensions/Robustness

C.1 Allowing for Elastic Order Quantities

Instead of assuming exogenous, perfectly inelastic quantities, we now add an additional stage after the static bargaining game in which buyers choose order quantities as a function of prices resulting from the auction. The modified timing of the game is:

1. Buyers simultaneously negotiate with all available manufacturers. In this process, sellers observe their private idiosyncratic cost shocks $\omega_{j,igt}$ and then submit bids. As before, the seller with the lowest bid wins and receives a price equal to the second lowest bid.

2. Order prices are realized.

3. I_{gt} chip buyers take the realized negotiated chip price as given, make production decisions, and choose the quantity of semiconductor chips needed as inputs for their downstream products (conditional on supplier) according to the following individual demand function:

$$\ln q_{igt}^{j} = \beta_{0,gt} - \beta_1 \ln p_{igt}^{j} + \boldsymbol{X}_{igt}\boldsymbol{\beta_2} \tag{C.1}$$

   where $\boldsymbol{X}_{igt}$ is a vector of observable buyer chip characteristics, $\boldsymbol{\beta}_{g,t}$ is a vector of parameters that can vary by generation and time period, and p_{igt}^{j} is the price the winning supplier receives.

4. Orders are filled and buyers and sellers make profits.

We assume order prices are fixed after the auction is run. This is consistent with either there being no renogotiation or bargaining being costless to the buyer such that the winning firm cannot extract more surplus from the buyer post auction.

A Equilibrium Bids

In a model with endogenous individual order quantities, firm j's expected per-unit payoff from buyer i as a function of its own bid b_j is:

$$\Pi_{ijgt} = \int_{b_j}^{\infty} \left(p_{igt}^{(2)} - \int AC(q_{igt}^{j}) dG_{q_{igt}^{j}}(p_{igt}^{(2)})\right) dF^{(2)}(p_{igt}^{(2)}) \tag{C.2}$$

where $p_{igt}^{(2)}$ is the price received (i.e., the second highest bid or second order statistic), $AC(q_{igt}^{j})$ is firm j's average cost at order quantity q_{gt}^{j}, $G_{q_{igt}^{j}}(p^{(2)})$ is the distribution function of potential

order quantities conditional on price received, and $F^{(2)}(p^{(2)})$ is the distribution function of potential second highest bids (i.e., distribution of the second order statistic).

A supplier's equilibrium bid b_j^* maximizes its per-unit payoff and is a function of average (variable) costs, which may take on a different value from marginal costs. Under constant marginal costs (as in our main specification) that are a function of total installed seller *capacity* ($f(\boldsymbol{q_{gt}^j}, Q_{gt}^j) = \zeta \cdot ln(Q_{gt}^j), \zeta \geq 0$), marginal cost equals average cost, and sellers bid their marginal costs in equilibrium. As a result, when order quantity is perfectly inelastic, this extension is equivalent to our main specification.

B Microfounding Individual Demand:

Equation C.1 is consistent with a Cobb-Douglas production function for a price-taking chip buyer:

$$Q_{igt}^D = q_{igt}^{\alpha}(M_{igt}^{other})^{\beta} exp(\epsilon_{igt}^D)$$

where Q_{igt}^D is downstream output, M_{igt}^{other} are other inputs to the downstream output, including other materials, labor, and capital, and $exp(\epsilon_{igt}^D)$ is a productivity shock. $0 < \alpha, \beta < 1$. The chip buyer's downstream profit function for a particular chip order when choosing chip quantities is:

$$\pi_{ijgt}^D = p_{igt}^D Q_{igt}^D - p_{igt}^j q_{igt}^j - p^{other} M_{igt}^{other}$$

where p_{igt}^D is the price of the downstream good, p_{igt}^j is its realized chip contract price, and $p^{other} M_{igt}^{other}$ is a price index reflecting the cost of other inputs (aggregated, for example, through some type of CES function).

Since the firm is price-taking, the first-order condition with respect to q_{igt}^j yields:

$$\ln q_{igt}^j = \frac{1}{\alpha - 1} \ln p_{igt}^j - \frac{\beta}{\alpha - 1} \ln M_{igt}^{other} - \frac{1}{\alpha - 1} p_{igt}^D - \frac{1}{\alpha - 1} \ln \alpha$$

which maps to an estimatable equation (Equation C.1) via $\beta_1 = -\frac{1}{\alpha - 1}$, and

$$\beta_{0,gt} + \boldsymbol{X}_{igt}\boldsymbol{\beta_2} = -\frac{\beta}{\alpha - 1} \ln M_{igt}^{other} - \frac{1}{\alpha - 1} p_{igt}^D - \frac{1}{\alpha - 1} \ln \alpha$$

C Estimation of Individual Demand Elasticity with Respect to Price

From the microfoundation in Appendix Section C.1B, our estimation equation is:

$$\ln q_{igt}^j = \beta_{0,gt} - \beta_1 \ln p_{igt}^j + \tilde{\boldsymbol{X}}_{\boldsymbol{igt}}\tilde{\boldsymbol{\beta}}_{\boldsymbol{2}} + \epsilon_{ijgt} \tag{C.3}$$

where $\tilde{\boldsymbol{X}}_{igt}$ are the observed controls for individual demand shifters and ϵ_{ijgt} includes the demand shifters unobserved by the econometrician.

We use GSA order-level data for this analysis (summary statistics are reported in Table 2). For each order, the dataset includes the quantity ordered, the unit price, and the manufacturing location. Our goal is to estimate the demand elasticity of quantity with respect to price. A key challenge is that observed prices reflect both demand and supply shocks, so OLS estimates are likely biased. To address this, we adopt an instrumental-variables approach: if we can find variables that shift local production costs (and hence prices) without directly shifting demand, we can use them to instrument for price and identify the demand response.

Our instrument choice is motivated by the location-specific resource intensity of semiconductor manufacturing and the potential for weather and natural disasters to disrupt production. Semiconductor fabrication requires substantial inputs such as water and stable operating conditions. Consistent with this, a 2023 survey of 100 senior decision makers in semiconductor companies reports that 73% cite natural resources (including water) as among the environmental factors posing the greatest risk to their business (James, 2024). Acute events such as heat waves and earthquakes can disrupt production in a particular manufacturing region, thereby affecting production costs and local supply conditions. At the same time, these shocks are unlikely to directly affect demand because semiconductor manufacturing is a highly international market in which chip buyers are often located in geographically diverse regions relative to chip manufacturers. For example, TSMC is located in Taiwan, but most of its clients—and the majority of its revenue—are located in other countries, such as the U.S. (around two-thirds of firm revenue in 2015) and China (around 10% of the firm's revenue in 2015).

We therefore compile a set of location-by-time measures of weather and natural-disaster intensity from multiple sources, including drought intensity from the SPEI Global Drought Monitor (Consejo Superior de Investigaciones Científicas (CSIC), 2023), disaster counts (e.g., droughts, earthquakes, wildfires) from Delforge et al. (2025), temperature and precipitation measures (World Bank), and earthquake magnitudes (USGS). We evaluate first-stage strength across a range of specifications. While drought intensity and earthquake counts can be strong in some cases (meeting the conventional $F > 10$ threshold), their strength is not robust to alternative clustering choices or sets of controls. By contrast, the Warm Spell Duration Index

(WSDI) and earthquake magnitudes are consistently strong and robust across the specifications we test. We therefore focus on results that use these two instruments.

In our baseline specification, we include manufacturing-location fixed effects and technology-generation fixed effects. Because manufacturers in different regions compete in the same global market for geographically diverse buyers, location fixed effects ensure that identification comes from within-location changes over time in weather/disaster intensity and their impact on costs and prices, rather than from cross-sectional differences across regions. Technology-generation fixed effects are important because advanced and mature nodes are effectively distinct markets with different buyers, applications, and different demand and supply conditions. We do not include quarter fixed effects, since doing so would absorb much of the time variation in weather/disaster exposure that provides identifying variation. We also do not add additional order-level characteristics (e.g., wafer size, number of layers) because the price and quantity variables used in this analysis have already been adjusted to account for these attributes.

The IV results are reported in Table C.1. Across specifications, the instruments are strong in the first stage, but the estimated demand elasticity is small and statistically indistinguishable from zero. One feature worth noting is that the first-stage relationship suggests that more intense heat spells and larger earthquakes are associated with lower observed prices, which may appear counterintuitive if one expects cost increases to raise prices. However, equilibrium prices are determined in a competitive environment where manufacturers and orders are matched endogenously. If adverse shocks reduce capacity or raise operational frictions at a location, the composition of observed transactions may shift toward orders that remain profitable to produce locally (e.g., those with particularly favorable terms, lower marginal costs, or stronger pre-existing contractual arrangements). Such selection and competitive adjustment can generate a negative reduced-form relationship between shock intensity and observed prices, even when the underlying shock increases production difficulty.

Overall, these estimates indicate that, conditional on manufacturing-location and technology-generation fixed effects, the elasticity of wafer quantities ordered with respect to price is small and statistically insignificant. This finding provides additional support for our baseline model specification, in which quantity demanded is perfectly inelastic in the short run and is driven primarily by downstream demand from electronics manufacturing rather than by short-run

fluctuations in semiconductor prices.

In an alternative specification, we omit technology-generation fixed effects, following Miao (2024), who estimates an elasticity between -1.488 and -1.686 using raw wafer input prices as an instrument. When we do not control for generation-specific differences in prices and demand, we obtain elasticity estimates that are similar in magnitude, as shown in Table C.2. The first-stage estimates are very similar to those in Table C.1, but the reduced-form relationship between log price and log quantity is much more positive, which mechanically implies a more negative IV elasticity estimate. This pattern is expected: both price and quantity are positively correlated with technology generation, so omitting generation fixed effects biases the reduced-form relationship between log price and log quantity upward.

Table C.1: IV Estimates of Demand Elasticity

	(1)	(2)	(3)
	Log quantity	Log quantity	Log quantity
Log price	-0.113	-0.133	-0.120
	(0.218)	(0.243)	(0.158)
Constant	7.556***	7.694***	7.606***
	(1.525)	(1.635)	(1.104)
Location Fixed Effects	Yes	Yes	Yes
Tech Generation Fixed Effects	Yes	Yes	Yes
Observations	11763	11763	11763
R-squared	0.144	0.146	0.145
First-stage Results			
WSDI	-0.034***		-0.034***
	(0.009)		(0.008)
Max Earthquake Magnitude		-0.046***	-0.049 ***
		(0.008)	(0.009)
First-stage F-value	14.71	38.61	22.21

Standard errors are clustered at the quarter level and reported in parentheses. Clustering at the location level instead yields substantially larger first-stage F-statistics. Prices (in $\$/m^2$) and quantities (in m^2) are defined as in Table 2 and are in logs. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.2: IV Estimates of Demand Elasticity (Without Technology-Generation Controls)

	(1)	(2)	(3)
	Log quantity	Log quantity	Log quantity
Log price	-1.095***	-1.211***	-1.138***
	(0.243)	(0.233)	(0.166)
Constant	14.289***	15.076***	14.583***
	(1.657)	(1.571)	(1.132)
Location Fixed Effects	Yes	Yes	Yes
Tech Generation Fixed Effects	No	No	No
Observations	11763	11763	11763
R-squared	0.112	0.114	0.113
First-stage Results			
WSDI	-0.031***		-0.031***
	(0.008)		(0.007)
Max Earthquake Magnitude		-0.046***	-0.045 ***
		(0.008)	(0.008)
First-stage F-value	14.72	34.98	22.06

Standard errors are clustered at the quarter level and reported in parentheses. Clustering at the location level instead yields substantially larger first-stage F-statistics. Prices (in $\$/m^2$) and quantities (in m^2) are defined as in Table 2 and are in logs. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C.2 Marginal Cost as a Function of Equilibrium Capacity Utilization

Not including utilization in our firm marginal cost function means that our counterfactual calculations may overstate policy effects. For example, when simulating how much U.S.-based manufacturers benefit from tariffs, the model could reallocate more demand to U.S. production than U.S. capacity can realistically accommodate. In practice, however, U.S. manufacturers can benefit from tariffs only to the extent that they have sufficient available installed capacity to serve the relocated demand.

Finding equilibrium bid strategies when marginal cost is a function of utilization and not just installed capacity is non-trivial theoretically and computationally. Having the marginal cost be a function of utilization makes a given seller's strategies across all buyers to which it submits bids interdependent, creating a much harder problem than a standard procurement auction. Below we expand on the theoretical and computational challenges in computing the full equilibrium with capacity constraints and discuss how we test the robustness of our results to binding capacity constraints.

A Theoretical Discussion

Firm j's expected per-unit payoff from buyer i as a function of its own bid for buyer i's contract b_i^j is:

$$\Pi_{ijgt} = \int_{b_j^i}^{\infty} \left(p_{igt}^{(2)} - AC(q_{gt}^j(p_{igt}^{(2)}, \boldsymbol{b_{-i}^j}))\right) dF^{(2)}(p_{igt}^{(2)}) \tag{C.4}$$

where $p_{igt}^{(2)}$ is the price received (i.e., the second highest bid or second order statistic). $q_{gt}^j(\boldsymbol{b^j})$ is firm j's total order quantity as a function of all of the bids its places $\boldsymbol{b^j} = \{b_1^j, \ldots, b_i^j, \ldots, b_{I_{gt}}^j\}$. $AC(q_{gt}^j(\boldsymbol{b^j}))$ is firm j's average cost for a given total quantity. $F^{(2)}(p^{(2)})$ is the distribution function of potential second highest bids (i.e., distribution of the second order statistic). For simplicity and given our individual order elasticity estimates, we assume in this discussion that order quantities are perfectly inelastic and exogenously set.

It is no longer an equilibrium strategy for firms to bid their marginal cost as in our main specification because firm j's bid for buyer i's contract affects its payoff to bidding for buyer k's $(k \neq i)$ contract. To see this more formally, note that seller j's total profits are $\Pi_{jgt} = \sum_i \Pi_{ijgt}$.

The b_i^j that maximizes expected profits satisfies the following first-order condition,as $dF^{(2)}(p_{igt}^{(2)}) = f^{(2)}(p_{igt}^{(2)})dp_{igt}^{(2)}$:

$$\frac{\partial \Pi_{jgt}}{\partial b_i^j} = \underbrace{-\left(b_i^j - AC(q_{gt}^j(\boldsymbol{b^j}))\right) \times f^{(2)}(b_i^j)}_{\text{(A)}} + \underbrace{\sum_{k \neq i} \int_{b_k^j}^{\infty} \frac{\partial AC}{\partial q_{gt}^j} \frac{\partial q_{gt}^j}{\partial b_i^j} dF^{(2)}(p_{kgt}^{(2)})}_{(B)} = 0 \tag{C.5}$$

This equation also illustrates how the first-order condition pinning down seller j's optimal bid for buyer i's contract also depends on the bids it places for all other buyers. Under our main specification, in which marginal costs are constant, part (B) equals zero, and it is optimal for each seller to bid their average variable cost, which equals marginal cost. Part (B) reflects how a seller internalizes the impact of auction i on the average costs of all of the other contracts for which it is bidding.

Solving this system of equations across multiple sellers and multiple buyers is nontrivial. The first challenge is to numerically approximate the distribution of the second-order price statistic for each buyer auction, which is complicated by the fact that our sellers face heterogeneous marginal costs. The second challenge is the dimensionality of the problem. If there are 10 sellers and 100 buyers, we need to solve for an equilibrium in terms of 1000 strategic variables.

While, a fixed point to the above equation does exist, we do not have a characterization of uniqueness. As such, we hope to leave solving this interdependent auctions equilibrium to future work.

B Robustness of Counterfactual Results to a 110% Capacity Utilization Cap

As a robustness check, we re-simulate all of our counterfactual scenarios while imposing a hard cap on capacity utilization at 110%, the approximate maximum utilization observed in this industry. Within each technology-generation and quarter market, we simulate buyer auctions sequentially: once fulfilling an additional order would raise a manufacturer's total contracted quantity above 110% of its installed capacity available for contract manufacturing, that manufacturer is removed from bidding for the remaining buyers in that market. If no manufacturer can accommodate an order within the cap, the order goes unfilled. We repeat this procedure for each of the 50 sets of bootstrap parameter estimates described in the notes of Table 5, for every scenario in Tables 6 and 7.

Table C.3 reports the results. Simulated profit levels are mechanically lower than in our main counterfactuals—for example, baseline profit falls from \$74.0B to \$56.9B—because the cap forces some orders to go unfilled rather than being absorbed by high-capacity manufacturers. However, the quantities of interest for our analysis, the differences between locating Fab 14 in Taiwan versus the U.S., are essentially unchanged: −\$1.7B versus −\$1.8B in the baseline, and +\$5.5B versus +\$5.9B when tariffs are added to the current regime. Our conclusions are therefore robust to imposing binding capacity constraints.

Table C.3: Simulated TSMC Profits with Capacity Utilization Capped at 110%

Policy scenario	Fab 14 in Taiwan	Fab 14 relocated to the U.S.	Difference
	(1)	(2)	(3)
No policy (baseline)	\$56.9B	\$55.2B	−\$1.7B
	[54.6,59.8]	[52.8,58.3]	[−2.3,−1.0]
U.S. export controls	\$55.5B	\$54.0B	−\$1.6B
	[53.2,58.4]	[51.6,57.1]	[−2.2,−0.9]
CN tax exemptions	\$54.9B	\$53.3B	−\$1.6B
	[52.6,57.9]	[51.0,56.3]	[−2.3,−0.9]
Export controls + CN tax exemptions	\$53.6B	\$52.1B	−\$1.5B
	[51.2,56.6]	[49.8,55.1]	[−2.2,−0.8]
U.S. import tariffs + export controls	\$55.9B	\$61.3B	+\$5.5B
+ CN tax exemptions	[53.3,59.2]	[58.7,64.7]	[4.7,6.2]

Notes: Each cell reports the simulated undiscounted sum of TSMC profits from 2004Q2 to 2015Q3, re-simulated under a hard 110% capacity utilization cap. Reported values are means across the 50 sets of $\widehat{\boldsymbol{\theta}}$ estimates described in the notes of Table 5; 90% confidence intervals, in square brackets, are based on the 5th and 95th percentiles of these simulations. Policy scenarios are calibrated exactly as in Tables 6 and 7.

C.3 An Upgrades Counterfactual: Fab 14 on a Lag-One Upgrade Schedule

In the observed data, TSMC upgraded Fab 14 generation by generation until it reached 40nm in 2008 and then stopped, while the industry frontier advanced by a further three to four generations through the end of our sample period. Our main counterfactuals hold this observed upgrade path fixed, so the losses from a U.S. location partly reflect a facility that is allowed to age. This appendix examines a counterfactual in which Fab 14 instead remains one generation behind TSMC's frontier throughout the sample period. The lag-one schedule reproduces Fab 14's observed upgrades through 2008 and adds three further upgrades, to 28/32nm, 20/22nm, and 14/16nm, in 2012Q4, 2013Q4, and 2015Q4. We re-run the full bootstrap simulation under this schedule for both fab locations and all policy scenarios, and we use our investment cost function, as in Appendix E, to predict the costs of the additional upgrade projects.

Table C.4 reports TSMC's simulated profits under the lag-one schedule. Continued upgrading shrinks the penalty from a U.S. location but does not eliminate it. Without additional policies, the difference between locating Fab 14 in the U.S. and in Taiwan narrows from \$1.8 billion in our main results to \$1.3 billion, and the differences under the policy scenarios narrow similarly. The tariff scenario again reverses the comparison, with a gain of \$5.3 billion compared with \$5.9 billion in our main results; the gain is smaller because a continuously up-

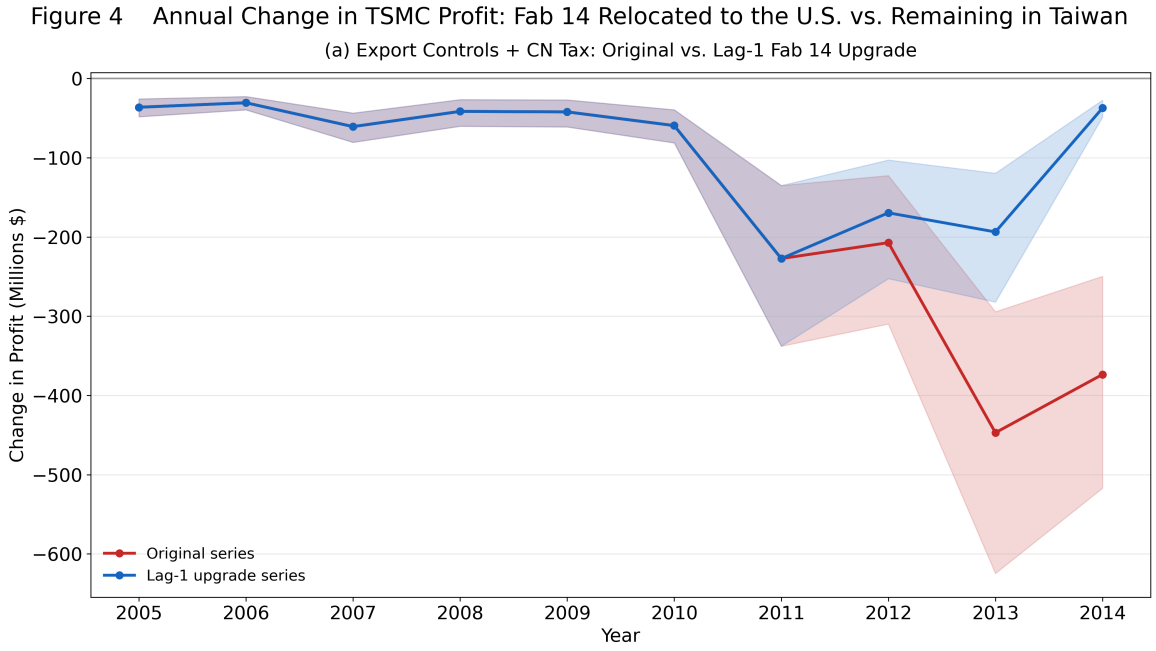
graded facility depends less on the mature, price-sensitive demand that tariffs redirect toward U.S. producers. These patterns are consistent with the maturation mechanism documented in Section 2.3: keeping the facility closer to the frontier reduces its exposure to entry-driven margin compression, which is the margin on which a U.S. location loses most. Figure C.1 shows this timing directly: the annual losses under the two upgrade paths coincide through 2011 and diverge once the synthetic upgrades begin, with the upgraded facility's late-year losses largely eliminated by 2014.

Table C.4: Simulated TSMC Profits with Fab 14 on a Lag-One Upgrade Schedule

Policy scenario	Fab 14 in Taiwan	Fab 14 relocated to the U.S.	Difference
No policy (baseline)	\$73.8B	\$72.6B	−\$1.3B
	[70.2,78.4]	[69.0,77.3]	[−1.9,−0.7]
U.S. export controls	\$67.8B	\$66.5B	−\$1.2B
	[64.6,72.2]	[63.4,71.1]	[−1.9,−0.6]
CN tax exemptions	\$70.5B	\$69.2B	−\$1.3B
	[66.5,75.4]	[65.2,74.0]	[−1.9,−0.7]
Export controls + CN tax exemptions	\$64.9B	\$63.7B	−\$1.2B
	[61.4,69.6]	[60.2,68.2]	[−1.8,−0.6]
U.S. import tariffs + export controls	\$64.9B	\$70.1B	+\$5.3B
+ CN tax exemptions	[61.3,69.7]	[66.7,75.0]	[4.6,5.9]

Notes: Each cell reports the simulated undiscounted sum of TSMC profits from 2004Q2 to 2015Q3 when Fab 14 follows the lag-one upgrade schedule described in the text. Reported values are means across the 50 sets of $\widehat{\boldsymbol{\theta}}$ estimates described in the notes of Table 5; 90% confidence intervals, in square brackets, are based on the 5th and 95th percentiles of these simulations. Policy scenarios are calibrated exactly as in Tables 6 and 7.

Fig. C.1: Annual Change in TSMC Profit from U.S. Relocation: Observed vs. Lag-One Upgrade Path



Notes: The figure plots the annual change in TSMC profit from relocating Fab 14 to the U.S. relative to remaining in Taiwan, under U.S. export controls and Chinese tax exemptions, for the observed upgrade path and for the lag-one upgrade schedule. Shaded bands are 90% confidence intervals across the 50 bootstrap estimate sets described in the notes of Table 5. The two series coincide until the first synthetic upgrade at the end of 2012 and diverge thereafter.

Table C.5 reports the predicted investment costs. The lag-one schedule is expensive: predicted investment for a U.S. Fab 14 rises from $7.8 billion under the observed upgrade path to $11.2 billion, an additional $3.4 billion of upgrade investment, and the premium over an equivalent Taiwanese facility widens from $1.1 billion to $1.4 billion.

The relevant comparison is between this incremental spending and what the firm receives for it. Because both instruments are proportional to investment, the larger project also draws larger incentives: at a 10% direct subsidy and a 25% ITC, the additional $3.4 billion generates about $1.2 billion of additional support, of which $0.9 billion is ITC, since upgrade projects place new qualified property in service. At the same time, the new schedule improves the simulated profitability of U.S. relocation from $1.8 billion below the no-relocation benchmark to $1.3 billion below it, a gain of $0.5 billion. Additional support and additional in-window profit together cover about half of the incremental spending, leaving roughly $1.7 billion uncovered.

Two things could close that gap. The first is a substantially larger credit. Holding the direct subsidy at 10% and the in-window profit gain at $0.5 billion, the ITC would have to supply about $2.6 billion of the additional $3.4 billion, implying a credit rate of nearly 75% on

the incremental investment. The 2025 increase from 25% to 35% supplies about $1.2 billion and still leaves $1.4 billion uncovered. The second is time. The lag-one upgrades arrive near the end of our sample, so most of the returns they generate fall outside the window, and the shortfall would have to be made up by operating profits earned beyond it. The asymmetry our main results identify is sharper on this margin: instruments tied to investment scale up automatically when the firm invests more, while the returns that would justify the additional investment accrue through continued operation, which those instruments do not reach.

Table C.5: Predicted Investment Costs of Fab 14 Under Alternative Upgrade Paths

Scenario	Projects	Investment	90% CI
Observed upgrade path, Taiwan	6	$6.7B	[5.8, 7.5]
Observed upgrade path, U.S.	6	$7.8B	[6.7, 8.8]
Lag-one schedule, Taiwan	9	$9.8B	[8.7, 11.2]
Lag-one schedule, U.S.	9	$11.2B	[9.7, 12.7]

Notes: Each row reports the predicted undiscounted sum of construction and equipment investment in Fab 14 from 2000 to 2015 under the indicated location and upgrade path. The lag-one schedule adds three upgrade projects to the six observed ones. Reported values are mean predictions across 30 independently trained MLP neural network models using different random seeds, with 90% confidence intervals based on the 5th and 95th percentiles of these predictions, as in Table 6.

This exercise quantifies the sense in which the ITC's coverage of upgrade investment does not resolve the timing mismatch. The credit covers one quarter of the upgrade cost by construction, and the upgrades themselves recover only part of the location penalty. At least within our simulation window, sustaining a continuously upgraded U.S. facility would therefore require support beyond the current incentive package even on the upgrade margin.

D Data Appendices

Table D.1: Data sources and variables

Dataset	Key Variables	Frequency & Coverage	Source
Wafer orders	Order-level quantity, price per wafer, fab location, manufacturing technology, number of layers; no information about buyer and manufacturer identity	Quarterly, 2003Q4-2015Q3; 2015Q2 data missing; worldwide representative sample of 11,927 orders	GSA
Fab capacity and investment	Fab-quarter-level installed capacity, manufacturing technology, construction & equipment investment, ownership, and product types; fixed fab characteristics such as location	Quarterly, 1995-2015; universe of fabs	SEMI
Contract manufacturing revenue	Firm-year-level contract manufacturing revenue in USD	Yearly, 2000-2017; data available for 246 firm-years; covers top pure-play foundries for 2000-2004 and top pure-play foundries and IDMs for 2005-2017	IC Insights
Contract manufacturing market revenue	Yearly overall revenue from the entire contract manufacturing market, broken down by pure-play foundries and IDMs, in USD billions	Yearly for 2003–2008 and 2014–2019	IC Insights
Various firm performance metrics	Firm-quarter-level contract manufacturing revenue, gross margin, and capacity utilization	Quarterly; data available for 277 firm-quarters between 2004 and 2015; covers only publicly listed pure-play foundries, no data for IDMs; data coverage varies by firm (see Appendix D.4)	Company reports
Technology revenue share	Firm-quarter-level revenue shares from different generations of manufacturing technology (e.g. "90nm," "40-45nm")	Quarterly; data available for 200 firm-quarters between 2004 and 2015; covers only TSMC, UMC, SMIC, Vanguard, and Hua Hong; data coverage varies by firm (see Appendix D.5)	Company reports
Region revenue share	Firm-quarter-level revenue shares from different geographical regions (e.g. "North America," "China")	Quarterly; data available for 152 firm-quarters between 2004 and 2015; covers only TSMC, UMC, SMIC, and Hua Hong; data coverage varies by firm (see Appendix D.6)	Company reports
Macroeconomic semiconductor demand indicator	Aggregated semiconductor sales based on net billings between semiconductor firms and their distributors and end customers	Monthly, 1986-present; one time series covering worldwide revenue	WSTS

D.1 Wafer Orders

The raw wafer orders data we obtained from the GSA contains 13,679 individual responses to their quarterly Wafer Fabrication & Back-End Pricing Survey for 2004–2015. We only retained observations where the wafers were manufactured using the CMOS process and where the wafer quantity and size were reported, resulting in the 11,927 observations reported in the main text.

D.2 Fab Capacity and Investment

The raw data we obtained from SEMI contains 2,265 unique manufacturing-related facilities. However, not all of them were relevant for our study. For example, because SEMI also serves the manufacturing supply chains for the display and photovoltaic industries, a substantial number of facilities in their database were used for the manufecturing of LED panels and other optoelectronic devices. To construct the sample of fabs for our study, we only keep the facilities that were used for commercial, volume production of semiconductors through the CMOS process. To accomplish that, we drop the following facilities:

1. Facilities that were not equipped with the CMOS process; most of those facilities manufactured optoelectronic devices (1,215 facilities remaining)

2. Facilities that did not perform commercial, volume production; most of those facilties were R&D or pilot facilities at research institutes or firms (1,005 facilities remaining)

The dataset has several different variables that capture the type of product produced at each fab, ranging from a coarse categorization (e.g. "Memory," "Logic") to the actual products (e.g. "DRAM," "MPU"). Whenever a fab was a dedicated foundry for outside orders, its product type at the coarsest level is labeled as "Foundry." We differentiate foundry capacity from IDM capacity using this label. We do not differentiate foundries that produce a wide range of products from specialty foundries that focus on a smaller range of products for two reasons. First, data quality for the finer product categorization is much poorer than that for the coarsest categorization and has lots of missing values. Second, our wafer orders data does not contain information about whether the wafers were manufactured at a specialty foundry or not, which forbid us from segmenting the foundry market further into different specialties.

For investment types, the raw data records many categories (such as changing wafer size, adding new lines, etc.). We classify these into five categories: new construction, tech upgrade, capacity expansion, upgrade and expansion, and maintenance. New construction occurs when there was no pre-existing facility (capacity equals zero) and a new facility must be constructed and equipped. A tech upgrade happens when the technology generation increases after the investment but the capacity is not expanded. Capacity expansion is when the capacity increases after the investment, but the technology generation is not upgraded. Upgrade and expansion

result in both increased technology generation and capacity at the facility. Maintenance occurs when both technology generation and capacity remain unchanged after the investment (i.e., there is no depreciation).

D.3 Yearly Contract Manufacturing Revenue

Information on yearly contract manufacturing revenue on the firm level is from publicly available data tables released by IC Insights, a semiconductor industry intelligence firm. Data from 2000 and 2001 is available from LaPedus (2003). Data from 2002 is available from LaPedus (2003) and LaPedus (2005). Data from 2003 and 2004 is available from LaPedus (2005). Data from 2005 to 2006 is available from Mutschler (2008). Data from 2007 is available from Mutschler (2008) and Chakraborty (2010). Data from 2008 and 2009 is available from Chakraborty (2010) and LaPedus (2011). Data from 2010 is available from LaPedus (2011) and Silicon Semiconductor (2012). Data from 2011 and 2012 is available from Silicon Semiconductor (2012) and Design & Reuse (2014). Data from 2013 is available from Design & Reuse (2014) and Bower (2016). Data from 2014 is available from Bower (2016). Data from 2015 is available from Bower (2016) and Design & Reuse (2018). Data from 2016 and 2017 is available from Design & Reuse (2018). When multiple data tables present duplicate observations for the same firm-year, the revenue numbers sometimes do not agree with each other potentially due to corrections and updates. In those cases, we arbitrarily select one observation per firm-year. The final dataset contains 246 firm-year observations between 2000 and 2017.

Information on yearly contract manufacturing revenue for the entire contract manufacturing market, including a breakdown by pure-play foundries and IDMs, is sourced from publicly available data figures released by IC Insights. Data for 2003–2008 is available from McGrath (2009), while data for 2014–2019 is available from Design & Reuse (2020b).

D.4 Quarterly Firm Performance Metrics

Data availability

Data on quarterly contract manufacturing revenue, gross margin, R&D spending, and capacity utilization are only available for publicly listed pure-play foundries that release their financial reports. Public firms that are IDMs do not separately list metrics from contract manufacturing only so we do not collect data from their financial reports.

TSMC: Company quarterly reports are available from 1997Q4 to present. Data on quarterly

revenue, gross margin, and R&D spending are available from 1997Q4 to present. Data on capacity utilization is available from 1996Q1 to 2005Q4. Reporting of capacity utilization stopped in 2006Q1, but wafer shipment and installed capacity are reported from 2005Q1 to 2015Q1.

UMC: Company quarterly reports are available from 2000Q1 to present. Data on quarterly revenue and gross margin are available from 2000Q1 to present. Data on R&D spending is available from 2000Q3 to present. Data on capacity utilization is available from 2001Q4 to present.

SMIC: Company quarterly reports are available from 2004Q1 to present. Data on quarterly revenue, gross margin, and R&D spending are available from 2003Q1 to present. Data on capacity utilization is available from 2003Q1 to present.

Vanguard International Semiconductor Corporation (Vanguard): Company quarterly reports are available from 2003Q3 to present. Data on quarterly revenue, gross margin, and R&D spending are available from 2002Q4 to present. Data on capacity utilization is available from 2002Q1 to 2015Q1.

Powerchip Semiconductor Manufacturing Corporation (Powerchip): Company semi-annual reports are available from 2000H1 to 2004H2 and from 2014H1 to present. Company quarterly reports are available from 2005Q1 to 2011Q2. Data on semi-annual revenue, gross margin, and R&D spending are available from 1999H1 to 2000H2 and 2002H1 to 2004H2. Data on quarterly revenue, gross margin, and R&D spending are available from 2005Q1 to 2011Q2. Data on capacity utilization is not available.

Tower Semiconductor Ltd. (Tower): Company quarterly reports are available from 2005Q1 to present. Data on quarterly revenue, gross margin, and R&D spending are available from 2004Q4 to present. Data on capacity utilization is not available.

Hua Hong Semiconductor Limited (Hua Hong): Company quarterly reports are available from 2014Q3 to present. Data on quarterly revenue and gross margin are available from 2013Q3 to present. Data on R&D spending is available for the years 2013 and semi-annually from 2014H1 to 2018H2. Data on capacity utilization is available for the year 2013 and for quarters from 2014Q1 to present.

Additional Data Cleaning

For revenue and R&D spending reported on the semi-annual level, we convert those results to quarterly level by equally splitting the revenue and R&D spending among the two quarters. For gross margin and utilization reported on the semi-annual level, we convert those results to quarterly level by assuming the quarterly results are the same as the semi-annual results.

Quarters where TSMC only reports wafer shipment and installed capacity, we calculate capacity utilization by dividing wafer shipment with installed capacity. Between 2005Q1 and 2005Q4, capacity utilization reported and capacity utilization based on our calculation are very similar.

After the additional data cleaning, we have 277 firm-quarter observations of revenue, gross margin, and R&D spending between 2004 and 2015. We have 196 firm-quarter observations of capacity utilization between 2004 and 2015.

D.5 Technology Revenue Share

Data on the revenue shares of different manufacturing technologies is available from company quarterly reports for the following firms and quarters:

- TSMC: 1998Q4 to present
- UMC: 1999Q3 to present
- SMIC: 2003Q1 to present
- Vanguard: 2002Q4 to present
- Hua Hong: 2014Q1 to present

We keep the 1,213 observations between 2004 and 2015 that belong to 200 unique firm-quarters.

D.6 Region Revenue Share

Data on the revenue shares of different geographical regions is available from company quarterly reports for the following firms and quarters:

- TSMC: 1998Q4 to present
- UMC: 1999Q3 to present
- SMIC: 2003Q1 to present

- Hua Hong: 2014Q1 to present

We keep the 634 observations between 2004 and 2015 that belong to 152 unique firm-quarters. In addition, we also have information about top pure-play foundries' share of revenue from China on the yearly leve for the years 2015, 2016, 2018, and 2019 from Design & Reuse (2017, 2020a). In 2015, China represents 8% of TSMC's revenue, 48% of SMIC's revenue, 10% of UMC's revenue, 5% of GlobalFoundries' revenue, and 55% of Hua Hong's revenue. In 2019, China represents 20% of TSMC's revenue, 59% of SMIC's revenue, 18% of UMC's revenue, 8% of GlobalFoundries' revenue, and 59% of Hua Hong's revenue.

D.7 Other Macroeconomic Data

WSTS's worldwide semiconductor billings data is publicly available and can be downloaded at `https://www.wsts.org/67/Historical-Billings-Report` as of February 5, 2024. The data contains four additional time series for revenue by end customer location (Americas, Europe, Japan, and Asia Pacific respectively). Given that our paper focuses on modeling and estimating chip designers' preferences for contract manufacturing location and contract manufacturers' location preferences, we do not make use of the information about end customer location.

We convert dollar amounts from other years into 2015 dollars using the consumer price index as in U.S. Bureau of Labor Statistics (2024). We convert currency amounts from Chinese Yuan Renminbi to US Dollars using Board of Governors of the Federal Reserve System (US) (2024a). Similarly, we convert amounts from Taiwan Dollars to US Dollars using Board of Governors of the Federal Reserve System (US) (2024b).

D.8 Determination of Technology Generations

Technology nodes in our data sources are reported in nanometers (nm) or micrometers (μm). We assign integer values $(1, 2, 3, ...)$ to represent technology generations, with higher values indicating more advanced technology. Each increment represents the next generation. We define these generations based on Moore's Law, which asserts that the number of transistors that can fit onto a microchip of the same size doubles approximately every two years. Consequently, the size of each transistor halves with each new generation. Historically, the technology node name indicates the width of specific physical features of a transistor, allowing us to deduce the generation. Mathematically, each new generation has a node size $\sqrt{2}$ (approximately 1.414) times smaller than the previous one. Alternatively, expressed logarithmically, the log of the

node size decreases by $log(\sqrt{2})$ (approximately 0.347) with each generation advancement.

Node names carried this exact physical meaning until the 28nm generation, which entered commercial production around 2012. After 28nm, shrinking the integrated circuits became increasingly challenging, and manufacturers continued to find new ways to improve the manufacturing process to deliver more powerful chips (Semiconductor Engineering, n.d.). Node names of newer generations still follow the numeric convention, but the numbers have lost their connection to the physical features on the chip. The best way to understand the meaning of a new node in recent years is as an umbrella term, which manufacturers use in the hope that the number can help to capture the set of new manufacturing technologies capable of delivering substantive improvements in the performance of the chips as compared to the previous generation (Hruska, 2021). Since around 2017, which falls outside of our sample period, node names have become even more inconsistent: technical specifications of chips manufactured using nodes with the same number no longer match from one manufacturer to the next, and node names are now often considered a marketing term. Because newer node names continue to follow the numeric convention, the generation assignment below applies throughout our sample period.

Let technology nodes greater than 4.6 μm be classified as technology generation 1. Following this pattern, technology nodes greater than 3.3 μm are considered technology generation 2, nodes greater than 2.3 μm are tech generation 3, and so on, until nodes greater than 0.0045 μm or 4.5 nm, which are classified as generation 22. Given a technology node of x μm, we can determine its corresponding tech generation g by solving the following inequality:

$$log(0.0045) + (21 - g)log(\sqrt{2}) < log(x) \leq log(0.0045) + (22 - g)log(\sqrt{2}).$$

E Predicting Investment Costs from Characteristics of Capacity Investment Projects

We fit multilayer perceptron neural networks to predict the investment amount for each project (in millions of dollars). We use the following project characteristics to predict this outcome:

- Manufacturer indicator variables (if a manufacturer accounts for less than 1% of the investment projects in our data, we replace the manufacturer's name with "Rare")
- Manufacturer's headquarter location indicator variables (China, North America, ROW,

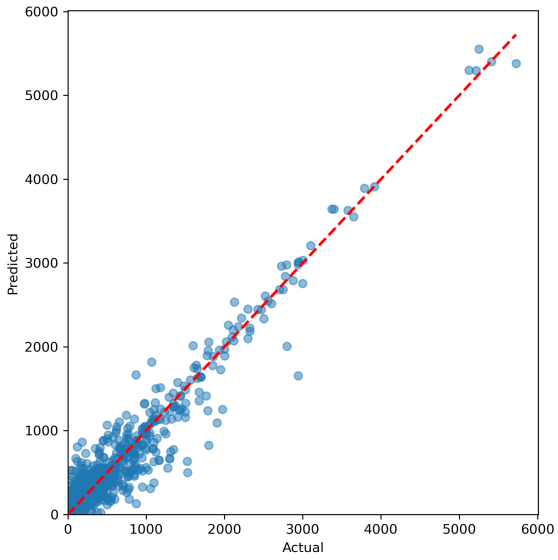
or Taiwan)

- Fab location indicator variables (China, North America, ROW, or Taiwan)
- Indicator for whether the fab is a foundry
- A continuous variable for the quarter in which the decision was made
- Log-transformed 200mm-equivalent capacity of the fab at the time of decision (with 1 added to the original value to avoid taking the log of zero)
- Log-transformed 200mm-equivalent capacity of the fab at project completion (with 1 added to the original value to avoid taking the log of zero)
- Fab's technology generation at the time of decision
- Fab's technology generation at project completion
- Fab's distance to the technology frontier at the time of decision
- Project construction duration
- Project equipment and ramp-up duration

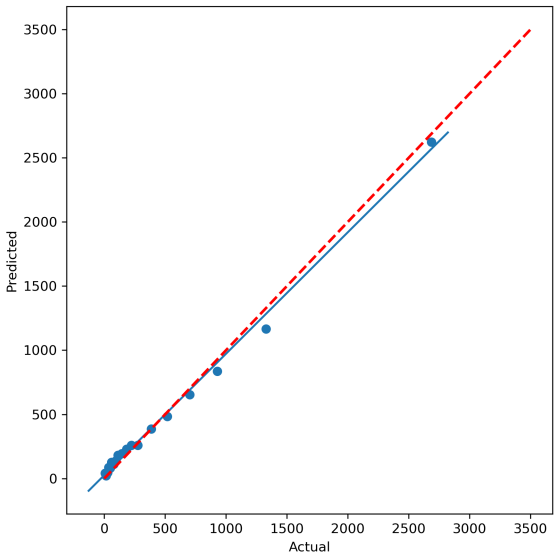
We transform the target variable, the project investment amount, by logging it and use the logged variable as the y-variable for model fitting. The loss function used is the mean squared error between the original variable against the inversely transformed out-of-sample prediction under five-fold cross-validation. The best layer size, learning rate, batch size, epochs, and activation function are determined through a grid search. The best configuration is: layer sizes = (2048, 32), learning rate = 0.00005, batch size = 8, epochs = 100, and activation function = ReLU. Below, we present the scatter plot of actual versus predicted values, a binned scatter plot of the same, and other in-sample and out-of-sample model fit metrics under the optimal parameter configuration.

Fig. E.1: Actual values and in-sample predicted values for investment amount

(A) Scatterplot



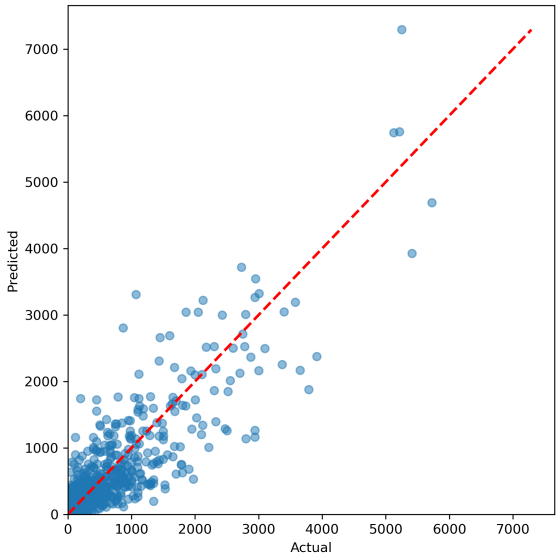
(B) Binned scatterplot



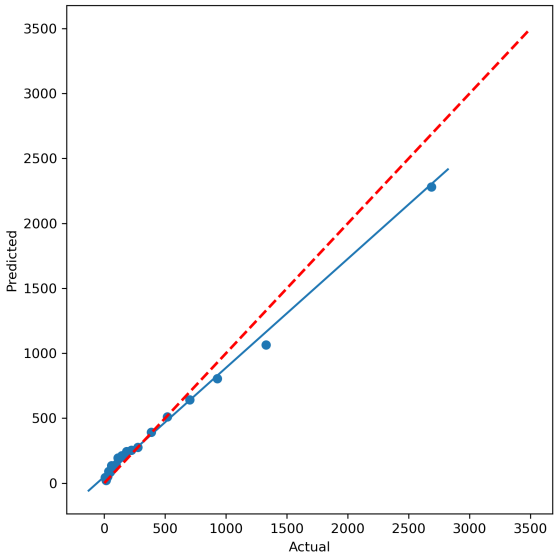
In-sample metrics: Mean Absolute Error (MAE): 104.7615, Mean Squared Error (MSE): 30369.7801, Root Mean Squared Error (RMSE): 174.2693 R^2: 0.9321.

Fig. E.2: Actual values and out-of-sample predicted values for investment amount

(A) Scatterplot



(B) Binned scatterplot



Out-of-sample metrics: Mean Absolute Error (MAE): 179.1811, Mean Squared Error (MSE): 108958.1110, Root Mean Squared Error (RMSE): 330.0880, R^2: 0.7565

Additional References for the Appendices

Board of Governors of the Federal Reserve System (US). Chinese Yuan Renminbi to U.S.

Dollar spot exchange rate [DEXCHUS], February 2024a. URL https://fred.stlouisfed.org/series/DEXCHUS. Retrieved from FRED, Federal Reserve Bank of St. Louis, February 5, 2024.

Board of Governors of the Federal Reserve System (US). Taiwan Dollars to U.S. Dollar spot exchange rate [DEXTAUS], February 2024b. URL https://fred.stlouisfed.org/series/DEXTAUS. Retrieved from FRED, Federal Reserve Bank of St. Louis, Februrary 5, 2024.

Nat Bower. Top 13 IC foundries by foundry sales in 2015. *Electronic Specifier*, 2016. URL https://www.electronicspecifier.com/news/analysis/top-13-ic-foundries-by-foundry-sales-in-2015. Accessed on January 8, 2024.

Pradeep Chakraborty. TSMC leads 2009 foundry rankings; GlobalFoundries top challenger! 2010. URL https://pradeeppoint.wordpress.com/2010/01/30/tsmc-tops-2009-foundry-rankings-globalfoundries-top-challenger/. Accessed on January 28, 2024.

Consejo Superior de Investigaciones Científicas (CSIC). SPEI global drought monitor. https://spei.csic.es/, 2023. Standardized Precipitation–Evapotranspiration Index (SPEI). Accessed January 2026.

Damien Delforge, Valentin Wathelet, Regina Below, Cinzia Lanfredi Sofia, Margo Tonnelier, Joris AF van Loenhout, and Niko Speybroeck. EM-DAT: The emergency events database. *International Journal of Disaster Risk Reduction*, page 105509, 2025.

Design & Reuse. Top 13 foundries account for 91% of total foundry sales in 2013. 2014. URL https://www.design-reuse.com/news/33775/major-2013-ic-foundries.html. Accessed on January 28, 2024.

Design & Reuse. Pure-play foundries boosting their presence in China. 2017. URL https://www.design-reuse.com/news/42827/pure-play-foundries-boosting-their-presence-in-china.html. Accessed on February 5, 2024.

Design & Reuse. China forecast to represent 22% of the foundry market in 2020. 2020a. URL

https://www.design-reuse.com/news/48819/china-foundry-market-in-2020.html. Accessed on February 5, 2024.

Design & Reuse. Pure-play foundry market on pace for strongest growth since 2014. 2020b. URL https://www.design-reuse.com/news/48686/2014-2024-foundry-sales-forecast.html. Accessed on January 1, 2025.

Kirsten James. The water challenge for semiconductor manufacturing: What needs to be done? *World Economic Forum*, July 2024. URL https://www.weforum.org/stories/2024/07/the-water-challenge-for-semiconductor-manufacturing-and-big-tech-what-needs-to-be-dor Accessed on January 17, 2026.

Mark LaPedus. IC Insights' pure-play foundry rankings for 1H03. *Electronic Design News (EDN)*, August 2003. URL https://www.edn.com/ic-insights-pure-play-foundry-rankings-for-1h03/. Accessed on February 5, 2024.

Mark LaPedus. China gains in 2004 pure-play foundry rankings. *EE Times*, 2005. URL https://www.eetimes.com/china-gains-in-2004-pure-play-foundry-rankings/. Accessed on January 28, 2024.

Mark LaPedus. Samsung lags in foundry rankings. *EE Times*, 2011. URL https://www.eetimes.com/samsung-lags-in-foundry-rankings/. Accessed on January 28, 2024.

Dylan McGrath. Forecaster predicts robust growth in foundry sales. *EE Times*, 2009. URL https://www.eetimes.com/forecaster-predicts-robust-growth-in-foundry-sales/. Accessed on January 1, 2025.

Ann Steffora Mutschler. Pure-play foundries comprise 84% of market, IC Insights says. *Electronic Design News (EDN)*, May 2008. URL https://www.edn.com/pure-play-foundries-comprise-84-of-market-ic-insights-says/#:~:text=The%20%E2%80%9CBig%204%E2%80%9D%20foundries%20%E2%80%93,sales%20of%20almost%20%2410%20billion. Accessed on January 8, 2024.

Silicon Semiconductor. Apple pushes Samsung's foundry sales yet again! 2012.

URL https://siliconsemiconductor.net/article/75826/Apple_Pushes_Samsungs_Foundry_Sales_Yet_Again. Accessed on January 28, 2024.

U.S. Bureau of Labor Statistics. Consumer price index for all urban consumers: All items in U.S. city average [CPIAUCSL]. Retrieved from FRED, Federal Reserve Bank of St. Louis, Februrary 5, 2024, February 2024. URL https://fred.stlouisfed.org/series/CPIAUCSL.